INTERNATIONAL COUNCIL FOR SCIENCE

INTERGOVERNMENTAL OCEANOGRAPHIC COMMISSION

WORLD METEOROLOGICAL ORGANIZATION

WORLD CLIMATE RESEARCH PROGRAMME

Report of the 11th Session of the JSC/CLIVAR Working Group on Coupled Modelling (WGCM)

3-5 September 2007
Hamburg, Germany

May

ICPO Publication Series No. 132

WCRP Informal Report No.8/2008

CLIVAR is a component of the World Climate Research Programme (WCRP). WCRP is sponsored by the World Meteorological Organisation, the International Council for Science and the Intergovernmental Oceanographic Commission of UNESCO. The scientific planning and development of CLIVAR is under the guidance of the JSC Scientific Steering Group for CLIVAR assisted by the CLIVAR International Project Office. The Joint Scientific Committee (JSC) is the main body of WMO-ICSU-IOC formulating overall WCRP scientific concepts.

Bibliographic Citation

INTERNATIONAL CLIVAR PROJECT OFFICE, 2008: Report of the 11th Session of the JSC/CLIVAR Working Group on Coupled Modelling (WGCM). May. International CLIVAR Project Office, CLIVAR Publication Series No. 132. (not peer reviewed).

CONTENTS

Action Items and Recommendations

1. Determine the JSC's response to re-branding WGCM as WCRP/CLIVAR WGCM-ACC
2. WGCM should exert some influence on the scientific agenda for the Modelling Summit.
3. Distribute the executive summary of the 2007 WGNE workshop to WGCM (K. E. Taylor to A. Pirani)

- WGCM recommends that WGOMD should comment on:The implications of the increasing resolution in atmospheric models that could overtake ocean resolution.
- Whether atmospheric resolution needs to match ocean resolution to better resolve the fine scale ocean structure?
- Whether fine scale structure is lost through the use of bulk formulae?
- How realistic is the CORE-III (freshwater perturbation experiment) forcing?

4. PCMDI will continue to support and update CMIP3
5. Modelling centres to be queried on whether there is willingness to open their CMIP3 data to less monitored public use.
6. WGCM confirms that PCMDI is the primary CMIP portal, reaffirming its leading role in WGCM/CMIP data distribution, together with modelling centres advocating data efforts in their own centres/countries.
7. CCCma to produce an ozone tropospheric-stratospheric dataset for use in the next round of coordinated experiments (G. Flato)
8. Follow up on which groups will be running the short term simulations, at what resolution, what ensemble size, with what initialisation and which will include interactive chemistry
9. WGCM endorsed CFMIP2 plans and recommends the use of the ISCCP-CloudSat-CALIPSO simulator (CICCS) in CMIP4 experiments. It also supports the storage of additional 3D model output, whilst recognising that this needs to be prioritised. WGCM also endorsed hosting CFMIP2 data together with the CMIP4 archive.
10. WCRP needs to communicate to users the need to quantify uncertainties in regionalisation (relative to other uncertainties, eg related to climate change). The peer-review process does not seem to be enough to filter out bad practice in the scientific literature. Consider putting regionalisation guidelines on the WGCM webpage.
11. Maintain links to ice sheet modelling community through J. Gregory, as well as maintaining interactions with CliC.
12. GEWEX to prepare list of cloud metrics for WGCM (S. Bony).
13. Inform WGCM of Climate-system Historical Forecast Project (CHFP) protocol once webpage is up to date (A. Pirani).
14. Gather response from modelling groups on interest in participating in Climate-system Historical Forecast Project (CHFP) once webpage is ready to be distributed.
15. Contact Johan Fedema to bring historical land use up at Amsterdam scenarios meeting and provide feedback to GCP if useful (G. Meehl).
16. Place scenario strategy onto WGCM webpage (G. Meehl to A. Pirani).
17. See list of chemical species prepared from modelling centre consultation. Note that ozone is not on the list. CCCma has stated that they will supply ozone dataset. Give list to K. E. Taylor to take to Amsterdam meeting (G. Meehl).
18. Link to atmospheric chemistry activities by passing chemical species list, response from modelling centres and CCCma O_3 plans to AC&C coordinators (P. Rasch, S. Doherty and A. R. Ravishankara) (G. Meehl).
19. After further iterations (C. Le Quéré, K. E. Taylor, J. Gregory) give details on final outcome of experimental design to assess carbon cycle feedbacks (to A. Pirani).
20. PCMDI to develop a metrics webpage (K. E. Taylor).

21. Put list of ice modelling activities from J. Gregory's presentation onto WGCM webpage (A. Pirani).
22. Communicate with SPARC community at the SPARC SSG held in Bremen, Germany on 18 to 21 September 2007 (M. Giorgetta).
23. Follow up with AC&C with regards to the inclusion of air quality in the short term experiments (G. Meehl).
24. Update recommendations for regional climate modelling at WCRP Modelling Summit (F. Giorgi).
25. Prepare a WGCM statement on the state-of-the-art of regional climate modelling (J. Mitchell).
26. Develop a WGCM webpage gathering information on regional climate modelling (A. Pirani).
27. Need to increase interactions with the wide community of stakeholders and end users of CMIP3 climate projections, to reduce likelihood of misinterpretation of model output and to identify specific needs that might be addressed in future CMIP phases. Initially consult with T. Busalacchi, C. Rosenflag (agricultural community) and M. Cane who have experience in this.
28. WCGM proposed members for a WGCM-WGSIP decadal prediction subgroup to expand on draft of decadal prediction experimental design plan as follows: subgroup to be headed by T. Stockdale and G. Hegerl with members J. Murphy, R. Stouffer, G. Meehl, M. Kimoto M. Giorgetta. The revised proposal should be sent to this group as well as WGCM in general. The JSC group assigned for decadal prediction also needs to be informed.
29. Explore the possibility of WGCM and WGSIP jointly holding a small decadal workshop.
30. Explore the possibility of holding a joint WGCM-AIMES meeting (G. Meehl).

1. Introduction

The 11th session of the JSC/CLIVAR Working Group on Coupled Modelling (WGCM) was generously hosted by the Max Planck Institute for Meteorology in Hamburg, Germany, on September 3-5 2007. The WGCM co-Chairs, J. Mitchell and G. Meehl welcomed the meeting participants, with particular appreciation for the presence of W.L. Gates, who was a founding member of the predessor of WGCM, the Steering Group on Global Coupled Models (SGGCM) when it was formed in 1990. J. Mitchell introduced the meeting and emphasised the need for quantitative climate science. The priority for WGCM is establishing the coordinated experiments using new baseline scenarios for a future AR5, while avoiding inhibiting climate science development through the need to produce new models and results inside the cycle time of the IPCC process.

The WGCM agenda has become very crowded in the last few years because of the ever increasing interest in climate change, and the development of more comprehensive climate models which include an increasing number of aspects of the full climate system. This year the form of the agenda was organized to allow a focus on the main priorities in improving simulations of climate and climate change and guide invited experts in how they can best help with those priorities. These priorities will also be communicated to the JSC, so that they can encourage and co-ordinate work across WCRP, especially in the projects, to improve climate modelling and the prediction of anthropogenic climate change in particular. This year the main items for discussion reflected the main points of emphasis of the WCRP Anthropogenic Climate Change (ACC) cross-cutting activity listed in Section 2.2, and this report is essentially structured to follow this list of items. The meeting agenda is given in Appendix I and the meeting participants are listed in Appendix II.

2. Overview of WCRP Activities of relevance to WGCM

2.1. JSC-XXVIII Session, Zanzibar, Tanzania (G. Flato)

WGCM will remain under the joint oversight of the JSC and the CLIVAR SSC. It was noted that the WGCM played a very valuable and visible role in the IPCC Working Group I Fourth Assessment Report through its coordination of the CMIP3 multi-model ensemble and organization of the analysis of model output. It was agreed that WCRP must focus on big science issues of relevance to society, and to focus on things that WCRP is uniquely positioned to do by virtue of its international connections and ability to foster broad collaboration. The JSC agreed that Core Projects and Working Groups remain central to WCRP.

ACTION: Determine the JSC's response to re-branding WGCM as WCRP/CLIVAR WGCM-ACC

Cross-cutting activities will serve to bridge gaps and improve coordination amongst projects and working groups. Some cross-cutting activities will be assigned to projects, some will remain under the auspices of JSC. The cross-cutting activity most relevant to WGCM is Anthropogenic Climate Change (ACC), described more fully below, in Section 2.2.

There was no clear consensus on the desirability of promoting a focus on a small number of very high resolution climate models and the associated supercomputing facilities. The WCRP Modelling Panel (WMP) will organize a 'Modelling Summit' in May 2008 to foster further discussion on WCRP's long-term modelling strategy and 'seamless prediction' (THORPEX has produced a white paper on seamless prediction). This will be the last activity for the WMP.

ACTION: WGCM should exert some influence on the scientific agenda for the Modelling Summit.

WCRP is experiencing serious financial difficulty that will affect all aspects of its operation. Efforts are underway to improve the situation. There is a need to continue to raise the visibility of WCRP and the work it does, particularly within the WMO and with policy makers at the national level. Projects and Working Groups should be active in this.

There is a need to address deficiencies in gender and geopolitical representation on WCRP committees. This must be considered seriously in future membership proposals.

2.2. Report on the WCRP Anthropogenic Climate Change (ACC) cross-cutting activity (G. Flato)

This cross-cutting activity will be overseen by the JSC, with contributions expected from all Projects and Working Groups. A JSC steering group has been constituted: H. Le Treut (Chair), J. Church, G. Flato, D. Griggs, and V. Ramaswamy.

The group met by teleconference, and in person (Paris, June 20, 2007) to discuss the overall objectives/priorities of ACC and to develop a work plan for the coming years. This will be fleshed out in the coming months. The intention is to build on existing efforts and show tangible output in the near term (e.g. concrete input to IPCC process, better connection to users), raising the visibility of WCRP as ACC is clearly a high-profile topic central to WCRP mandate.

Some activities that should be focused on:

- Quantitative projections of climate change: 21st Century and beyond, including facilitation of future multi-model ensembles and analysis thereof (already a WGCM core activity).
- Fostering improvements in climate models. e.g. clouds and aerosols, carbon cycle feedbacks (joint with AIMES), development of metrics.
- Forcing scenarios: balancing 'science' needs with those of impacts/policy community. Involves IPCC WGs II and III. More discussion at Netherlands meeting on New Scenarios in October 2007.
- Serving the 'impacts community' (WG-II). e.g. regional downscaling (can WGCM be more active in this?); decadal prediction (ocean initialisation, ensemble methods, etc. – joint with WGSIP, WGOMD, CLIVAR Panels).
- High impact or emerging issues. e.g. ice sheets (presumably CliC and PAGES); extreme events; coupling between air quality and climate change.

In terms of the impacts community (WG-II), there is an absence of a coherent community-representative group at the international level that WGCM can communicate with. More needs to be done by modelling activities in each country in terms of working with the impacts community locally. Scientifically, two questions need to be addressed:

- Determining the ability of current models at simulating regional variability, let alone regional climate change, and whether we can downscale from climate models.
- Quantifying uncertainty at regional levels

WGSIP has already addressed this by providing a statement, instead of taking a coordinating role, on what can be done and which regions would benefit from regional downscaling, including the fact that statistical downscaling is not successful in observation-poor regions.

2.3. WGNE: Report from the Workshop on Systematic Errors in Climate and NWP Models (K.E.Taylor)

The JSC/CAS Working Group on Numerical Experimentation (WGNE) and Program for Climate Model Diagnosis and Intercomparison (PCMDI) hosted a Workshop on Systematic Errors in Climate and NWP Models during the week of February 12, 2007. The principal goal of the workshop was to increase understanding of the errors and their causes in coupled climate models. The WGCM, GEWEX Model and Prediction Panel (GMPP), and the Working Group on Seasonal to Interannual Prediction (WGSIP) all contributed to the planning and coordination of the workshop. Over 180 scientists participated.

Although the meeting was basically organized around time-scales (short, intermediate and longer), the first day was devoted to invited presentations on cross-cutting overview talks (e.g., metrics, systematic errors in the tropics, parameterization errors) and specific phenomena (e.g., ENSO, land-surface and cloud interactions, radiation). Besides the presentations (both oral and poster), break-out sessions were convened each afternoon to focus on a range of specific topics including "perturbed physics ensembles," metrics, diurnal cycle, ENSO, tropical biases, and monsoons and intra-seasonal oscillations.

Among the general conclusions that emerged from the workshop discussions were the following:

- Climate models can benefit from being tested on short ("weather-forecast") time-scales when deviations from the observed weather trajectory can be more easily traced to specific causes.
- Likewise, seasonal prediction with climate models could reveal much about their ability to simulate intermediate time-scale phenomenon (few weeks to a year). We should expect to learn much about their ability to accurately represent the atmospheric water cycle, the partitioning of precipitation (convection vs. stratiform), and the vertical profile of heating. All of these are critical to capturing intra-seasonal variability (e.g., MJO).
- Examination of the diurnal cycle should be a future focus because it can reveal problems in the representation of precipitation processes, land-sea exchanges (e.g., sea breezes), propagating convective systems, and monsoons.
- The limitations to accurate simulation of weather in climate models resulting from the constraints on resolution imposed by practical considerations need to be more fully understood and quantified. Similarly, ocean model simulation of boundary currents and eddies require higher resolutions than is normally accessible within current computing resources. Thus, targeted simulations of climate with much higher resolution models are called for.
- Interest in establishing a suite of performance metrics for climate models was expressed as a way of encouraging quantitative assessments of the relative merits of different models, but there were objections to establishing any single index of performance without rigorous justification for its value.

ACTION: Distribute the executive summary of the 2007 WGNE workshop to WGCM (K. E. Taylor to A. Pirani)

2.4. Report from WGSIP of issues relevant to WGCM (T. Stockdale)

WCRP Task Force on Decadal Prediction (from JSC)

- A modest initial work plan has been proposed to the JSC involving three member (at least) ensembles of 20 years with two start dates (1965 and 1994), looking at the role played by initial conditions and forcing (1965 with 1994 forcing and vice versa). This is based on experience from the ENSEMBLES Project and is similar to what was done for Stream 1, while Stream 2 proposes more start dates.
- The JSC endorsed this proposal as a cross-cutting activity between WGCM and WGSIP. WGSIP has recently agreed to work on this, if we can do so jointly with WGCM. A subgroup of WGSIP will be formed to handle WGSIP's input.
- There is a question as to whether and how this work relates to the proposals for the "initial value" climate runs to 2030.

The S/I community is increasingly dealing with issues which are also relevant for the climate change problem

- "Global warming" is an important part of 6-month forecasts - even with the right initial conditions, temperatures after 4 months of integration are substantially dependent on eg CO_2 levels. The impact of volcanic aerosols is also visible.
- Despite careful initialisation and the short forecast range, models have clear failures, eg with regard to trends in NH summer SSTs over the last 25 years.
- Nonetheless, shorter range SST forecasts are generally good for dates a long way into the past - suggests upper/mixed layer initialisation is broadly OK for last 40-50 years.
- Analysis of forecast errors and initial condition uncertainties shows that model errors are by far our dominant problem.
- A good SST climate does not guarantee good forecast performance.
- Models are not yet able to predict the seasonal timescale reasonably well, and in particular the reliability of probabilistic forecasts is still very unsatisfactory. This suggests that at least the dynamical aspects of regional climate change may also be

unreliable with today's models (eg recent European summers ..).

- Resolution is an important question (T95, T159, T399, more?)

Testing of "climate change" models in seasonal forecast mode

- Many errors of coupled ocean-atmosphere models are visible in integrations of a few months. It is easier to diagnose these errors by running integrations from observed initial states and comparing with observed data valid for a specific period, than to compare a long run of a model with observations of "mean climate". This is particularly true for recent high quality datasets.
- A straightforward experimental protocol can be given, based on seasonal forecasting experience, which allows the upper ocean to be initialised, such that suitable ensemble integrations of up to a year can be made. The ocean is initialised with wind and SST data - ocean data assimilation is not needed.
- WGSIP are offering to produce recommendations on this for use by the model development and assessment communities, though without proposing to organize any formal intercomparisons. This simple protocol would involve moderate amounts of integration (eg for ENSO: 20 years * 4 dates/year * 3 members * 7 months = 140y, for teleconnections: 30 cases * 11 members * 4 months = 110y).
- As well as model biases, the actual forecast skill of models can also be assessed. ENSO forecast skill, in particular, does not need huge samples. There is a natural link here to the Climate-system Historical Forecast Project (CHFP, formerly TFSP) experimentation, which WGSIP hopes many of the models used by WGCM will participate in (baseline set of experiments is approx 600y of integration).
- At the WGSIP meeting in June 2007, the possibility of IPCC assessment of the seasonal forecast capability of AR5 models was raised. WGSIP intend to send a member to discussions on the scoping of AR5.

WGSIP have given much thought to data handling issues for the CHFP Experiment

- New CF proposals have been made (to allow multi-model and ensemble forecasts to be described), and these also should be suitable for climate use.
- The CHFP Experiment will use a distributed approach, initial software will be openDAP
- Request that the possibility of joint work is kept in mind, eg with CMOR and data servers (although WGSIP has very little ability to lever resources in this area).

ENSO in IPCC models (from E. Guilyardi)

- Need better procedures for assessing ENSO in IPCC-class models
- A two-pronged proposal involving (a) metrics and (b) simulations has been developed in discussion with various groups. Once the feedback is complete, this will be proposed formally as part of a US CLIVAR project.
- At this stage, comments on the proposal and perhaps an endorsement of the general idea would be helpful (the proposal is not yet sufficiently finalised to be endorsed as such). [Note that the 'simulations' part of the proposal maps onto the general framework of testing IPCC models in seasonal forecast mode, discussed above].
- WGCM might also consider what role they would like to play in such a project (WGSIP and the CLIVAR Pacific Panel are other natural players).

2.5. Report from WGOMD of issues relevant to WGCM (H. Banks)

The work of WGOMD will input mainly to the discussion on improving ocean climate models. The key items that WGOMD would like WGCM to note are:

- The Coordinated Ocean-ice Reference Experiment (CORE) design enables ocean-ice models to be compared in a relatively controlled setting by using common atmospheric data forcing provided by Large and Yeager (2004). Work with the repeating annual year CORE-I involves seven global ocean-ice models which have each been run for 500 years. A manuscript documenting the experimental design and model results is in preparation (Griffies et al, 2008, to be submitted to Ocean Modelling).

- The CORE design is now being used increasingly in climate modelling centres worldwide (e.g., GFDL, NCAR, MPI, CSIRO (plans), Hadley Centre (plans)). This development allows for a more controlled setting to compare the ocean-ice model components commonly used in fully coupled climate models.

- WGOMD has initiated the development of an online Repository for Evaluating Ocean Simulations (REOS). This repository will include metrics and methods for evaluating ocean-ice simulations. There are many efforts ongoing in the community to highlight opportunities for ocean modelers to evaluate their simulations based on comparison to observational data. REOS aims to provide a guide for modelers, as well as the opportunity for observational datasets to be featured and discussed from the modelling perspective.

- Various groups are moving forward with the development of coupled climate prediction systems for the purpose of decadal forecasts. This problem involves many issues of middle and high latitude oceanography previously largely ignored for the more mature tropical forecasting systems. WGOMD aims to play a role in the development of these systems, initially by providing guidance and communication between various groups.

ACTION WGCM recommends that WGOMD should comment on:
- The implications of the increasing resolution in atmospheric models that could overtake ocean resolution.
- Whether atmospheric resolution needs to match ocean resolution to better resolve the fine scale ocean structure?
- Whether fine scale structure is lost through the use of bulk formulae?
- How realistic is the CORE-III (freshwater perturbation experiment) forcing?

Despite the fact that unstructured grids are rendering analysis more difficult, WGOMD strongly recommends that vector fields should not be interpolated to regular grids, so the output archive should support both regular and native grid data. PCMDI endorses the fact that native grid storage is needed for ocean dynamics, though the volume of data involved is massive at higher resolution. Software is needed so that the wider community can deal with multiple grids.

3. Emissions scenarios and Coordinated Experiments (J. Meehl)

We are now on the threshold of including earth system model-type components (carbon cycle, chemistry, aerosols, dynamic vegetation) in “standard” global coupled climate models (atmosphere, ocean, land, and sea ice; AOGCM) used for climate change projections. This has led to a joint WGCM/AIMES effort (involving an AGCI session, August 2006; and a joint WGCM/AIMES meeting, September 2006), with the objectives of:
1. Identifying what new components are ready now or will be ready in the next six months for inclusion in AOGCMs.
2. Establishing communication through WCRP, IGBP and Scenarios Consortium to coordinate activities in preparation for climate change simulations that could be performed with this next generation of models (possibly for an IPCC AR5).
3. Proposing a strategy encompassing an experimental design for 21st climate change experiments with these models (near term and longer term time frames) to reduce uncertainties that involve the relevant processes.
4. Specifying the requirements for these new models in terms of time series of constituents from new emission scenarios (impacts/adaptation/mitigation, stabilisation).

3.1. Overview of CMIP Data Management and the legacy of the multi-model CMIP3 dataset (D. Bader, K. E. Taylor)

The success of the WCRP CMIP3 multi-model dataset (Meehl et al, 2006), which, through the coordination of the WGCM, was created by the modelling centers and collected and disseminated by the Program for Climate Model Diagnosis and Intercomparison (PCMDI, is apparent both in its scientific impact on the IPCC Fourth Assessment Report (AR4) and its ongoing contributions to the published literature. In Chapters 8-11 of the AR4, which rely heavily on climate model simulations and which cover model evaluation, understanding and attribution, global projections, and regional projections, over 70% of the figures are based on the CMIP3 results.

Perhaps more impressive is the fact that four of the seven figures in the AR4's Summary for Policy Makers were made possible by this dataset. The multi-model perspective afforded by CMIP3 provided a stronger scientific basis for assessing the reliability of the projections, which is important in evaluating the likely impacts of future climate change on, for example, ecosystems, agriculture, and society.

The standardization and distribution of the CMIP3 dataset has been the most labor intensive part of the process for modelling centers. The use of NetCDF and CF was essential for the success of the CMIP3 data dissemination and CF is now institutionalized through WCRP/ WGCM with support from Unidata, BADC and PCMDI.

The interest in CMIP3 from scientific researchers continues unabated. Over 1000 scientific subprojects have been registered to access the data, and each week typically a dozen or more new subprojects are registered. Over 300 journal articles have been published or have been submitted, and the rate appears not to have diminished over the last two years, as illustrated in Figure 1.

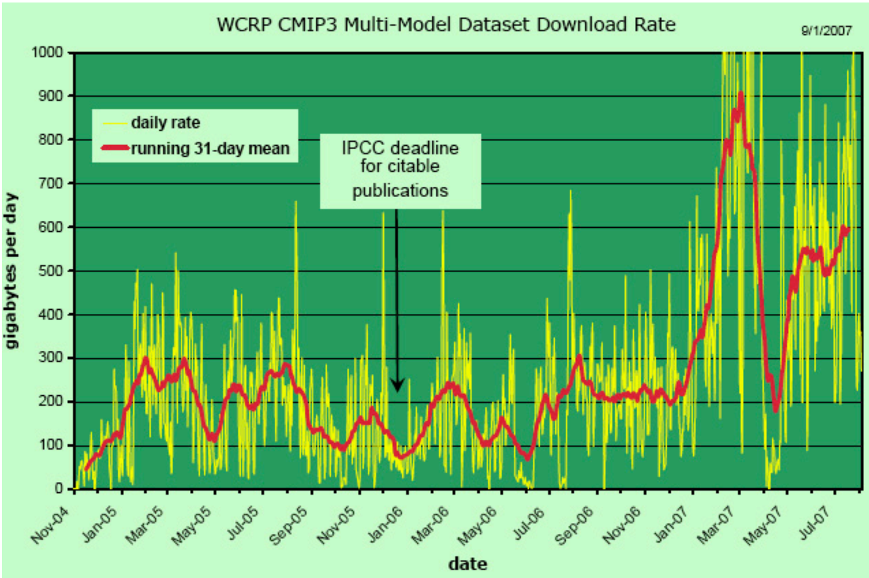


Figure 1: WCRP CMIP3 multi-model dataset download rate

The considerable interest in the CMIP3 data and the rather loose control over who currently can download that data raises the question as to whether the WGCM should encourage modelling groups to remove all restrictions on access to the data (i.e., "make it public"). This issue should be considered by the WGCM.

ACTION: PCMDI confirms it will continue to support and update CMIP3.

WGCM endorses PCMDI support of the CMIP3 dataset at least until the next intercomparison experiment (~2011), especially due to its continued growth in use. This includes finding errors, supporting modelling centres wanting to replace runs etc.

The current system of manual registration and cursory review by the CMIP panel is in place since some modelling groups do not allow their model output to have unlimited access. If the data archive were to be opened up completely (with the users agreeing to appropriate "terms of use"), then an automatic registration procedure could be implemented, reducing the burden on PCMDI and the CMIP panel. Only US data is currently being served without a password (through OpenDAP). If data is requested for commercial use, the user should be directed to contact individual centres to obtain permission to access the data.

ACTION: Modelling centres should be queried on whether there is willingness to open the CMIP3 data to public use.

Planning has begun among a group of centers, led by PCMDI, to accommodate model output from the next round of coordinated community coupled model experiments (including, optionally, representation of the carbon cycle). A database distributed among several sites (i.e., modelling centers and data distribution centers) will be linked together by the Earth System Grid (ESG), coordinated by the Global Organization for Earth System Science Portals (GO-ESSP). The ESG is designed to provide seamless access to all the data. Compared with CMIP3, the next round of simulations is expected to produce a much larger set of output (in the realm of hundreds of terabytes, a factor of 10 increase in complexity compared to CMIP3, with increased importance on metadata), and consequently it will be impractical (and for other reasons undesirable) to collect it all at a single location. Portals for the Grid will sit at several distributed sites, taking into account individual centre needs, security etc, but with a common interface.

PCMDI has partnered internationally with other groups, who provide complementary expertise, in order to assure that a service will be in place to fully serve the needs of the climate modelling community as it prepares for future assessment activities. PCMDI will provide centralised registration for data access and will still be able to act as a repository for centres that cannot store their output. The system needs to be vetted by the community to ensure that it is solid and reliable.

ACTION: WGCM confirms that PCMDI is the primary CMIP portal, reaffirming its leading role in WGCM/CMIP data distribution, together with modelling centres advocating data efforts in their own centres/countries.

WGCM thus confirms the unambiguous endorsement of the continuing leadership by PCMDI in the provision of CMIP data. Further along the planning process for CMIP4, individual modelling centres need to be contacted to determine the cost of the nodes, the development of the back ends of the portals to be tailored to individual centre security requirements etc, so they can make the necessary preparations and apply for funding. Aiming for a prototype data system by 2009, one or two qualified engineers will be required at each site to support this process.

3.2. Review of proposed experimental design for the next coordinated experiments addressing mitigation/adaptation scenarios (CMIP4) (J. Meehl)

In July 2006, participants of the Aspen Global Change Institute 2006 Session on Earth System Models (organized by G. Meehl, WGCM co-chair, and K. Hibbard, chief scientist, IGBP AIMES) proposed two classes of models and experiments to address two time frames and two sets of science questions for the next round of coordinated experiments for future climate change (Meehl and Hibbard, 2007, and summarized by Hibbard et al, EOS, 2007):

1. *Near-Term (2005-2030) "Decadal Prediction" Experimental Protocol*
   High resolution (at least 0.5°) simulations with no carbon cycle, some chemistry and aerosols, and a single mid-range scenario (since scenarios do not spread significantly until after mid-century), using an ensemble approach. Such projections would examine, for example, the likelihood of changes in extremes on the regional scale.

The following issues need to be addressed in terms of the near-term experimental design:

- Does coupled initialisation of the observed state matter, i.e. is there decadal predictability from an observed initial state that would improve projections for the 25-year time frame? Yes, in some regions, for some quantities, particularly over the oceans (Smith et al, 2007). Several efforts are underway, applying different coupled initialisation techniques, to examine the mechanisms that could generate decadal predictability (UKMO Hadley Centre, GFDL, NCAR, IFM-GEOMAR, INGV, etc.).

- What is signal to noise for climate changes on the regional space scale for the 25-year time scale? Would changes be detectible? Yes, in some regions and for some fields (e.g. temperature), with significant signals emerging on short (~30 year) timescales (Wu and Karoly, 2007).

- How important is time-evolving chemistry/aerosols for regional climate change? Is temporally evolving chemistry necessary, for example for ozone, or are time-slice experiments sufficient?

- Should more experiments be run with more ensemble members at a lower resolution or fewer ensemble members at a higher resolution? Based on future computing capabilities, the option of more ensemble members at a higher resolution will be possible.

- Is there a scientific case for higher atmospheric model resolution? Yes, e.g. WGNE recommends that higher resolution reduces systematic errors and gives a better resolution of extremes. Increasing the horizontal resolution increases the number of wintertime cyclones in the Northern Hemisphere close to observed numbers (Jung et al, 2006).

- The role of land use change will be examined by Integrated Assessment Modelling (IAM) groups supplying integrated time-evolving future land use input to use in climate models.

- In lieu of using full chemistry models, should specified stratospheric ozone be used? This could be a possibility for some groups and would involve the generation of time-evolving 3D ozone concentrations (could this also be done for aerosols). G. Flato will provide ozone dataset for use by modeling groups.

ACTION: Form a WGCM-WGSIP subgroup to develop a proposal for a near-term climate change coordinated experiment, to be evaluated within a year's time.

See Appendix III for a draft proposal for coordinated experimentation to study multi-decadal prediction and near-term climate change prepared by the newly forming WGSIP-WGCM subgroup. This experimental design will be updated and revised at an Aspen Global Change Institute session in June, 2008, with the product taken forward for final approval at the WGCM meeting in September, 2008.

2. *Long term (to 2100 and beyond) Experimental Protocol*
   Lower resolution (roughly 1.5°) simulations, using Earth System Model (ESM)-type capabilities that include the carbon cycle and specified or simple chemistry and aerosols, with representative concentration pathways (RCPs) quantify feedbacks in the climate system in response to different scenarios.

The Integrated Assessment Modelling (IAM) community in conjunction with WGCM and AIMES are proposing four RCPs from 2005-2300: an A2-type reference (no mitigation ~8Wm^{-2} at 2100), and three mitigation scenarios, low (~3 W^{-2}), medium (~4.5 Wm^{-2}) and high (~6 Wm^{-2}), to 2300.

Originally, scenarios were derived with a forward-looking approach from socio-economic calculations, as illustrated in Figure 2a. A "reverse approach" is now being proposed for scenarios with WG3 supplying a few RCPs derived from climate scenarios, WG1 determining the emissions and feeding them back to WG2 and WG3 to derive the socio-economic impacts, illustrated in Figure 2b.

a) Forward-looking approach

b) Reverse approach

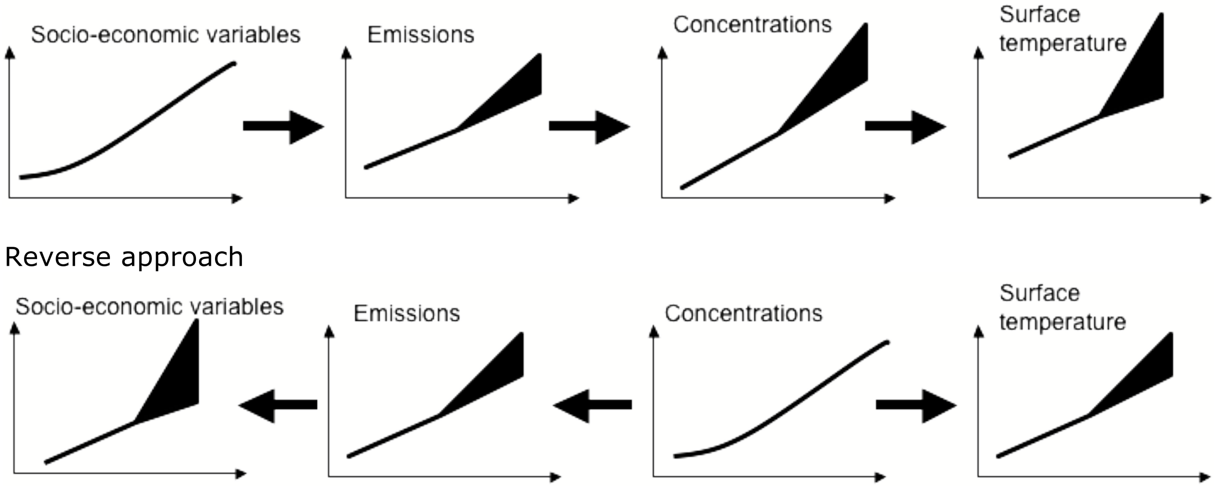


Figure 2: Schematic of traditional forward looking approach (a) of starting from socio-economic variables to derive emissions, concentrations, then temperature and other climate changes from climate models, and the new proposed methodology (b) where the starting point is concentrations run in climate models, that are used to derive emissions and then socio-economic factors that will achieve these emissions. The wedge associated to the graphs indicates the uncertainty associated with the process. The forward looking approach is characterized by uncertainty that grows in the direction of the response of the climate system, while the reverse approach the uncertainty grows in the direction of the socio-economic factors (from Meehl and Hibbard, 2007).

Long-term runs provide an opportunity to contribute to a policy perspective on avoiding consequences of climate change (e.g. mitigation/stabilisation/adaptation). New mitigation scenarios run with earth system models will have implicit policy actions to target future levels of climate change. However, since we can only mitigate part of the problem, and we will have to adapt to the remaining climate change, the challenge is to use climate models to quantify time-evolving regional climate changes that human societies will have to adapt to.

Concerns were also raised (G. Hegerl) that there is a problem with the non-CO_2 forcing in the runs with an interactive carbon cycle. Non-CO_2 should be included in order to provide meaningful projections and to be able to compare the 20th Century model results to observations. However, the idea of estimating the carbon cycle feedbacks based on comparing a coupled carbon cycle simulation with a carbon cycle under constant climate assumes no other forcings. EMICs with a coupled carbon cycle (as already used in AR4) may help with tests on how to deal with the non-CO_2 forcing, or idealized experiments suggested by K. Taylor.

RCPs will be turned over to WGCM finalised by about September, 2008, following feedback that was solicited from the modelling community and the IPCC Expert Meeting on New Scenarios held in Amsterdam in September 2007. Most modelling groups will have finalised the new model versions by 2009 and begin running experiments by 2009-2010.

In terms of the feedback solicited from the modelling community, the following issues have been raised so far:

- How should experiments be run with land use changes
- How to smoothly transition from 20th Century forcing to 21st Century scenarios
- What quantities should be stored?
- Which chemical constituents should be included in scenarios (eg HFCs, PFC, CH_4, etc)
- At what altitude should emissions be released?
- Should dust and sea salt be included?
- Should aerosol indirect effects be included?
- Should organic carbon from volatile organic compounds (VOCs) be included?
- The list of experiments is too long and needs to be prioritised
- Are the short-term experiments too short? Should they not be extended to 2050?

- How to include ozone and stratospheric processes
- Can permissible climate change range be produced for each land type?
- How can the difference likely to arise between prescribed CO_2 emissions from land use changes and CO_2 emissions calculated from the carbon cycle be standardized?
- Should the experimental design be tested using EMICs first?
- Should wild fire be taken into account?
- Will PCMDI be able to cope with archiving what is likely to be ten times the data produced by AR4?
- Should idealized calibration runs of natural with or without anthropogenic forcings be run for the 20^{th} Century?
- Should the simulations be coordinated with what is already being done by the ENSEMBLES project stream 2 simulations A1B with land use change for non-mitigated reference; "E1" mitigation scenario developed within ENSEMBLES (stabilised at 2100 near 475ppm CO_2 equivalent)?

These issues will be revisited at the WCRP Modelling Summit planned for May 2008, and at the next WGCM meeting in September, 2008.

3.3. Summary of proposed CMIP4 coordinated experiments (and proposed order of priority)

A. Standard calibration runs with AOGCM or ESM:

pre-industrial control (300 yrs)
1% CO_2 increase to doubling (100 yrs)
slab $2XCO_2$ for climate sensitivity

B. Long term

AOGCM and ESM Experiment 1 (specified concentrations)

a. 20th century (~1870-2005), 135 yrs, all-forcings
b. scenarios
 -8 Wm^{-2} A2-like, 100 yrs, encourage one member to 2300
 -3 Wm^{-2} low overshoot, 100 yrs, encourage one member to 2300
 -Optional in order: 4.5 Wm^{-2} stabilisation
 6.0 Wm^{-2} stabilisation

[5 member ensembles encouraged]

ESM Experiment 2 (fully coupled carbon cycle driven by emissions)

8 Wm^{-2} A2-like, 100 years, one member to 2300
3 Wm^{-2} low overshoot, 100 yrs, one member to 2300

Optional in order: 4.5 Wm^{-2} stabilisation
6.0 Wm^{-2} stabilisation
[5 member ensembles encouraged]

ESM Experiment 3 (specified concentrations, no climate change)

1% CO_2 (compare to TCR 1% run)
[1 member]

C. Short term

1965-2035, 70 years, single scenario (4wm^{-2} stabilisation)
[encourage higher resolution, and at least 10 member ensemble]

Emphasise:
-Extra ensemble members for short term climate change
-Possibly initialised from observations
-Possibly higher resolution than long term simulations

General Guidelines

- Run CFMIP simulator, perhaps though time slices
- Save 3-hourly data
- Run anthropogenic and natural forcings separately for the 20th Century
- Run short AMIP style runs to analyse feedback (see K. E. Taylor's proposal in Appendix IV)
- Collect atmospheric model output on native model vertical levels and archive ocean data on native grid as well as interpolating to regular grids.

J. Gregory commented that, in contrast to CMIP3, where mid-range scenarios were prioritised, extreme cases of no mitigation and full mitigation are prioritised here.

G. Hegerl also commented that International Detection and Attribution Group (IDAG) members (often from their IPCC experience) also have some suggestions about scenarios, coming from an impact perspective, eg a string of emission reduction scenarios and some business as usual to see what each option "buys". These may help WGII and III to do cost-benefit analyses of mitigation scenarios. Also, we need some consistent scenarios to be able to compare and connect results to AR4, since the media and policymakers seem very interested in comparing previous and new IPCCs.

3.4. Report from the Task Group on Data and Scenario Support for Impact and Climate Analysis (TGICA) (N. Nakićenović)

In April 2006, the IPCC decided that its role would be limited to "catalyzing" the development of scenarios in the scientific community and the decision to hold an expert meeting in Amsterdam in September 2007 was taken to (a) "catalyze" efforts in the scientific community; and (b) increase involvement of developing country/EIT participants.

There are three major scenario user communities:

- <u>Climate modelling community</u> – need scenarios to provide a coherent, internally consistent, time-path for Earth System Models.
- <u>Impacts modelling community</u> – need scenarios to provide a coherent, internally consistent, time-path to assess the consequences of potential climate changes and to set the context for adaptive strategies.
- <u>Integrated assessment modelling community</u> – to provide a coherent, internally consistent, emissions scenarios and to assess the costs of emissions mitigation

A more limited IPCC role leads to the question of what level of coordination is needed among different users and providers of scenario information? (international/ regional/national/local) and what mechanisms are need to be put in place to achieve this coordination, in the absence of an IPCC coordinating role? These issues need to be addressed since the integrated assessment modelling (IAM) community is not tightly organised and the Impacts-Adaptation-Vulnerability (IAV) community is even less coordinated (though TGCICA does contribute).

An international consortium has been established to facilitate the coordination of scenario development efforts led by the International Institute for Applied Systems Analysis (IIASA), the Energy Modelling Forum (EMF) Stanford University and the National Institute for Environmental Studies (NIES).

Representative Concentration Pathways

Many questions need to be answered regarding the RCPs:

- Role of baselines?
- What forcing agents? (detailed specification, including spatial and temporal resolution)
- Coordination of assumptions across IAM and ESM?
- Criteria and candidates? Selection process?

A major purpose of the session at Snowmass meeting in 2007 was to prepare an input on benchmark scenarios for discussion at the September IPCC expert meeting. This constitutes the preparatory phase of the four phases involved in the preparation of scenarios, illustrated in Figure 3 (the associated timeframe could be the basis for setting the timing for a possible IPCC

AR5 for the year 2014, and listed here:

1. The Preparatory Phase provides RCPs (to be completed by June-July 2008
   - Long-term, HIGH reference scenarios
   - Long-term, MID stabilisation scenario (for calibration)
   - Long-term, HIGH stabilisation scenario
   - Long-term, LOW stabilisation scenario

2. Phase 1 explores a broader range of socio-economic scenarios
   - Reference & stabilisation levels
   - Stabilisation accession
   - Societal development paths
   - Technology dynamics
   - Carbon cycle & climate
   - "Overshoot" stabilisation
   - Regional scenarios

3. Phase 2 - Integrated scenarios
   - Link ESM scenarios with global Phase 2 scenarios
   - Incorporation of net fluxes from ESM results to create partially consistent scenarios
   - Scalability of ESM results
   - Add to library for vulnerability, impacts, and adaptation research

4. Phase 3 – Iterative process to create consistent treatment of mitigation, impacts and adaptation in a new set of community integrated scenarios
   - Agriculture-land-use-terrestrial carbon cycle-ecosystems
   - Revised energy supply (e.g. hydro, biomass) and demand (e.g. heating/cooling)
   - Etc.

.

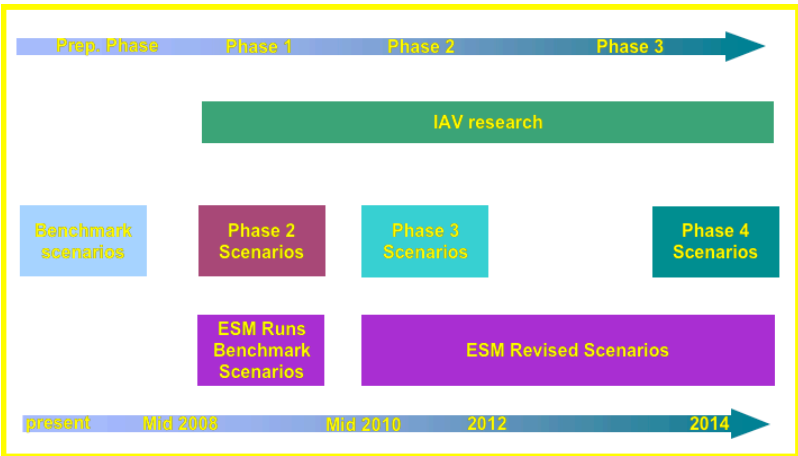


Figure 3: Relative timeframe of the four phases in generating scenarios

An additional high baseline reference scenario is proposed here in addition to the original initial stabilisation scenarios proposed in the Aspen design (Meehl and Hibbard, 2007):

- High, no stabilisation baseline reference – 8.5 W/m^2
- High stabilisation level – 6 W/m^2
- Medium stabilisation level – 4.5 W/m^2 (for high resolution, short term experiments), at a mid-point between high and low stabilisation, capturing stabilisation within category IV of Figure 4.
- Low stabilisation level – 3 W/m^2

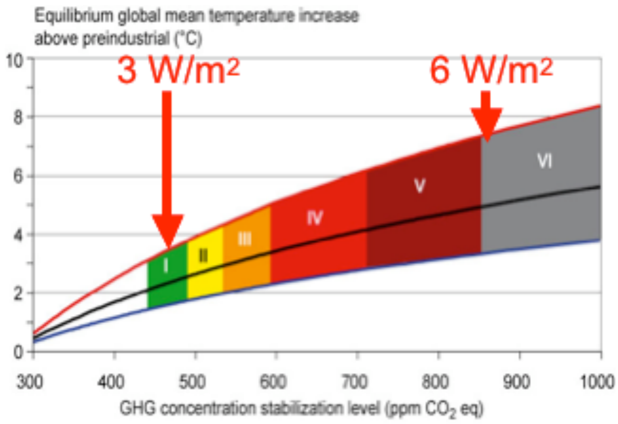


Figure 4: Long-term mitigation: GHG concentration stabilisation levels and equivalent equilibrium global mean temperatures

A question that needs to be addressed is what geographic distribution, and with what resolution, the forcing agents will be needed by ESMs:

- Well mixed gases: Geographic distribution may not matter.
- Reactive gases: Geographic distribution might matter, e.g. CH_4.
- Aerosols: Geographic distributions definitely matter, but at what geographic scale (e.g. ½° X ½° for up to 2030)?
- Land-use and land cover: Geographic distributions definitely matter, but what variables and what units? Land-use harmonization?

Criteria

The following criteria need to be met by the scenarios:

- Published in the peer-review literature as scenarios are needed by mid-2008)
- Roughly correspond to the levels prescribed
- Contain endogenous carbon cycle and atmospheric chemistry representations that reflect the present state of the art
- Provide all of the forcing agents at appropriate scale of disaggregation
    - This may require revisiting existing scenarios to obtain the appropriate information, e.g. agriculture, land use, and land cover.
- Have been vetted extensively
    - These scenarios will receive extensive scrutiny.
    - There is insufficient time to begin with new (unpublished) scenarios.
- Must be useful for Impacts-Adaptation-Vulnerability (IAV) community

Recommendations

1. The scenario details still need to be worked out with the ESM modelling community. That needs to happen very quickly.
2. The Consortium should ask it's members for an official expression of interest on the part of modelling teams to provide RCPs ahead of their Phase 2 contributions.
    - This includes a commitment to provide the resources needed to complete the task.
    - This also includes a demonstration that the modelling team can provide a scenario that meets ALL of the criteria outlined above.
3. The consortium should move forward to identify THREE long-term stabilisation scenarios from the existing peer-reviewed literature, defined around the steady-state radiative forcing levels listed above.

Potential additional work includes the possible extension of reference and stabilisation scenarios to the year 2300, for example an A2 scenario based on coal, though the question remains of how stylized such a scenario would be.

All gases need to be benchmarked, not just CO_2. The scenarios also need to take into account the possible carbon-climate feedback that occurs in 2005-2010 (C. Le Quéré**).** An ozone tropospheric-stratospheric evaluation should be solicited with the aim of producing it by June-July 2008.

ACTION: CCCma to produce an ozone tropospheric-stratospheric dataset for community use (G. Flato).

3.5. Recommendations from WCGM to IPCC Expert Meeting on New Scenarios (Amsterdam, 19-21 September 2007) (J. Mitchell)

The experiments should be prioritised according to the scenarios, starting with the high, no stabilisation scenario at 8.5 W/m^2 (where interesting feedbacks are expected) and the low stabilisation scenario at 3 W/m^2 (there is the possibility of negative emissions being derived), then the 4.5 W/m^2 and 6 W/m^2 stabilisation scenarios. For the high and low scenarios, an extra carbon experiment should be run. For high resolution experiments, recommend scenario 4.5 W/m^2, ideally together with a low resolution experiment and a carbon experiment.

A common initialisation strategy for high resolution short term scenarios has not been defined yet so a recommendation cannot be made yet to the international modelling and impacts communities and it may be premature to declare that these short term predictions will be provided to an AR5 at the meeting in Amsterdam.

However this problem is currently being addressed and there is a wide suite of experiments, depending on different scientific questions, that could be run (high v. low resolution, different initialisation methods with the same model, etc). Instead of confirming that short term predictions will be produced, this could be phrased to say that this question is currently being addressed in terms of assessing what the predictive capabilities are for periods of 20-30 years (T. Stockdale). Modelling groups will be performing these kinds of experiments anyway as the result of stakeholder demand in terms of understanding the impact of climate change on extremes, even with low resolution simulations. The short-term simulations are currently defined as being high resolution and initialised from observations, and so addressing extremes and predictability respectively. The issue of high resolution does not necessarily have to be a requirement.

Several groups are already working towards simulations initialised from observations (eg Met. Office Hadley Centre, GFDL, NCAR, INGV, CSIRO). It would be good to know how many groups will be running these short term simulations, at what resolution, with what initialisation and which will include interactive chemistry.

ACTION: Which groups will be running these short term simulations, at what resolution, what ensemble size, with what initialisation and which will include interactive chemistry?

3.6. Report from the Global Carbon Project (P. Friedlingstein)

The Global Carbon Project (GCP – www.globalcarbonproject.org) was established in 2001 in recognition of the enormous scientific challenge and fundamentally critical nature of the carbon cycle for Earth sustainability. The scientific goal of the project is to develop a complete picture of the global carbon cycle, including both its biophysical and human dimensions together with the interactions and feedbacks between them.

The following issues need to be addressed when designing the next set of coordinated experiments using the reverse approach, ie starting from concentrations:

- The climate-carbon cycle feedback will impact the "compatible" emissions, not the climate projection
- Large positive carbon feedback may be missed as the implication for negative emissions may lead to these being modified to be more realistic to derive economic calculations.

IPCC WG III may not accept emissions derived from prescribed CO_2 concentrations using the reverse approach, described in Figure 1b. In the alternative proposed by the GCP, emissions should be derived from socio-economic drivers and then used to derive CO_2 concentration, from which the climate change (ΔT) and the equivalent "compatible" emissions are derived used to quantify the associated socio-economic impacts.

According to the GCP, the order of priority of the experiments proposed by the Aspen White Paper (Meehl and Hibbard, 2007, WCRP tech note), described in Section 3.1. would be Experiments 1, 3, and 2.

3.7. Suggestions from the International Detection and Attribution Group (IDAG) on the design of 20th century simulations for next IPCC (G. Hegerl)

Start date for future simulations
Connect model simulations (particularly the high resolution model runs) to the 20th Century, with an ideal start date around 1950; the earlier the better particularly for attributing sea level rise. This would improve the signal-to-noise ratio in detecting and attributing changes in noisy variables such as extremes and rainfall that are expected to become very important in AR5, as well as for producing probabilistic forecasts of these important quantities.

Spin-up suggestions:

- From observed ocean data (pro: reduces spread and is like a decadal prediction. Concern: drift back to model climate and lack of separation between ensemble members, limiting our ability to attribute changes to forcing rather than ocean state; will also limit ensemble spread).
- From lower resolution runs (pro: ensemble size increases, good for attribution studies since only driving factor is forcing, no assumptions about ocean state. Lower resolution runs, start time 1850 to 1900, or before would be even better).

Multiple ensembles
For all forcing simulations, maybe a natural-only ensemble to about 2010 (with no future volcanoes and extrapolated solar forcing), and an ensemble that enables the separation of the greenhouse gas (GHG) contribution by either having GHGs alone, or all except GHGs. This is helpful for probabilistic future predictions. Large ensembles would be best and if a minimum ensemble size is suggested, it would be good to suggest multiple members. Larger ensembles, higher spatial resolution and shorter runs will increase the ability to quantitatively assess changes in extremes, ideally producing enough daily data (eg from a100-year run, or five 20-year ensemble members) to realistically estimate a 100-year return period.

M. Allen suggests that a fully-interactive ocean model may not be needed for short-term prediction or timeslice experiments and could be replaced using a trended slab ocean and varying SST patterns to add more uncertainty and represent different states. 100 member ensembles could be run to generate more weather, assess the importance of SST patterns and to look at the changing statistics of extreme events.

Control simulation
Enough control simulation needed to sort out drifting runs as has been done for AR4 analyses. It may be possible to use controls from the lower resolution models even if fingerprints from high-resolution are used, but as this is not optimal, some control of the high resolution models would be extremely helpful.

Forcing documentation
The forcing should be clearly specified and documented. The question of whether a consistent, synchronized forcing should be proposed raises the concern is that this would make spanning the uncertainty range much harder (by not having any forcing uncertainty in the ensemble) and that there is too much uncertainty in some forcing terms to propose a single realization. Having a range of forcings will address uncertainty in these terms. Forcing data needs to be delivered better than last time, and monthly (providing just a reference (and sometimes the wrong one as happened in AR4) is not sufficient).

Tuning
Some concern was expressed that some groups may be starting to use 20th Century warming

for model tuning. What is done to tune models should be clearly documented so that the models can be validated (what was looked at for model tuning (means and trends), and over what time period and temporal and spatial resolution).

Variables
The following data have been suggested in addition to the monthly and daily data collected for AR4 (http://www-pcmdi.llnl.gov/ipcc/standard_output.html#Experiments):

- Saving daily data for at least 50 yrs.
- Monthly mean Tmin and Tmax and surface humidity
- Daily mean specific surface humidity
- Daily averaged ocean surface heat budget components and surface windstress
- Daily snowcover and major river flows into oceans (alternatively, monthly mean surface and total runoff?). K. Taylor pointed out that CMIP3 collected monthly-mean surface runoff and total runoff including subsurface drainage – would that be sufficient?
- Possibly (see AR4) monthly mean freshwater flux to oceans.
- Archiving some impact relevant variables such as carbon fluxes, forest fire, land cover, ice sheets etc where available.

The following high-resolution, 6-hrly resolution data (would be useful for storm studies):

- daily max of winds at (1) 850mb U, V; more prudent may be winds at all levels from 850 to 250mb, and (2) lowest model level or 10-meter U,V (helps establish intensity and compare to observations – should a standard scheme be used to diagnose?)
- Sea level pressure
- Temperature at 300mb, 500mb, and a few levels in-between (in order to compute vertically averaged temperature in the 300-500mb layer, temperatures are needed to determine warm core and thus tropical storms – it may be possible to use lower temporal resolution for this).
- Geopotential height at 1000 and 200mb

Other variables used for storm detection include: 200-1000Hpa thickness, 850Hpa temperature, 200Hpa and 700Hpa temperature and 300Hpa u and v.

To reduce data size, one might save the more detailed height-dependent wind and temperature info and SLP every 24 hours (instantaneous values) but keep 6-hour resolution on a key wind variable for vortex tracking (850mb relative vorticity is most commonly used). Using daily max would also save space compared to saving 6-hourly data. The latter could probably be linked to a nearby warm core vortex without too much trouble even if the timing is not exact. It would be good to have lots of years (say minimum 40 yrs at end of perturbation and 40 yrs of control, 40 yrs recent observed period?; or multiple ensembles to, for example, circumnavigate the AMO. An optional10-yr slice of 3-D state of the atmosphere may be very helpful for nesting models.

3.8. Benchmark Experiments Proposed to Understand ESM response (K. E. Taylor)

The set of experiments, in the Aspen White Paper (Meehl and Hibbard, 2007) and discussed in Section 3.1. focus primarily on quantifying uncertainty in carbon cycle feedbacks, but largely omit consideration of other important feedbacks in the climate system (e.g., clouds, water vapor, sea ice). A modified suite of experiments has been subsequently proposed (*see Appendix 4: "Maximizing the Value of Community-Coordinated Experiments with AOGCMs and ESMs"*), which augments the so-called "realistic" scenarios proposed in the White Paper with idealized experiments in which only the concentration of carbon dioxide is prescribed to change. These idealized experiments, along with some short diagnostic simulations, will enable performance of a fully unified feedback analysis of coupled carbon-climate projections.

The analysis approach is to evaluate the top of the atmosphere radiative response by separating the radiative response to different forcings:

- "Fast" radiative responses" (i.e., radiative "forcing"), eg:
    - GHG immediate impact on TOA radiative fluxes
    - Stratospheric adjustment (on time-scales much shorter than time-scale of global temperature change)

- "Slow" radiative responses (i.e., "feedbacks"), eg,
  - Stefan-Boltzmann damping
  - Water vapor / lapse rate
  - Clouds
  - Surface albedo

The aim is also to diagnose carbon cycle feedbacks:

- "Fast" carbon cycle feedbacks (evident even in the absence of climate change), e.g.:
  - Ocean carbon uptake
  - "fertilization"
  - Net Primary Productivity (NPP) and effect on surface albedo, evapotranspiration (& water vapor & clouds) and on the heating profile of surface layers of the ocean
- "Slow" carbon cycle feedbacks (evident as climate & subsurface reservoirs of carbon change), e.g.:
  - Ocean circulation, temperature, & mixing
  - Dissolved carbon, ocean acidity, & buffering
  - Vegetation types
  - Plant decay (oxidation) rates
  - Plant responses to changing climatic conditions

Details of the proposed experiments and analyses are given in Appendix 4.

3.9. Contribution of paleoclimate to the improvement of climate modelling (P. Bracconot)

Paleoclimate contributes to a better understanding of the mechanisms of climate change, the identification of key feedbacks operating in the climate system, and the evaluation of the capability of climate models to reproduce climates different from the modern one. The paleoclimate community works on different time scales with a hierarchy of models. The studies using coupled GCM concern mainly :

1. Simulations of the last millennium:
   - There are now several 1000 yr long runs with IPPC class models that can be used to assess the natural variability of the climate system.
   - These simulations are also very useful for climate detection studies to better understand the role of natural forcings (volcanism, solar) and trace gases prior to the industrial area.
   - The expertise is distributed in different centres. Data syntheses are emerging that can be used to evaluate the results. At the moment the different initiatives are not coordinated at the international level.
   - The difficulty for model evaluation is the small signal to noise ratio and the non-synchronicity of the different records
   - WGCM should provide recommendations to promote model intercomparisons and the coordination of dataset syntheses

2. Simulations of key time periods in the past:
   - Efforts have been developed within the Paleoclimate Modelling Intercomparison Project (PMIP) since 1991
   - Database for simulations of mid-Holocene (6000 yrs BP) and LGM (21000 BP)
   - Simulations for other time periods with working groups to share the expertise
   - Better signal to noise ratio
   - Several data syntheses available

The propositions below only concern this second approach and the contribution from the PMIP project. These items will be discussed at the next PMIP2 committee in Exeter (end of September 2007). Several workshops are already planned for the coming year:

- Autumn 2008 : US (B. Otto-Bliesner)
- 2009 : Japan (A. Abe-Ouchi)
- 2011 : Berne in Switzerland during the next INQUA conference

Paleoclimate modelling could have interesting contributions to:

- Improving climate models
- Serving the impact community: the PMIP2 database is open and anybody can propose a subproject. Several projects concern impact studies
- High impact or emerging uncertainties.
- Some studies also consider regional modelling and down scaling, so the methodologies and some of the results may be worth considering.
- Process modelling : use process models to understand either mechanisms or environmental records (vegetation, isotopes, ocean biology).

Several approaches are currently developed:

1. Test periods with well documented data and ensembles of model simulations (= PMIP for mid-Holocene and LGM)
   - Development of benchmark diagnostics
   - Revision and improvement of already existing datasets
   - New diagnostics and methodologies will be developed
2. Understanding role of different feedbacks including (clouds, ocean, vegetation, snow and ice)
   - Need to develop cross projects between PMIP and CMIP3 (already done) and other MIPs
   - Have more models included in PMIP2 database
3. Test and comparison of interactive vegetation models - these models are now developed in several groups. Paleoclimates offer the opportunity to test the interactive vegetation models.
4. Analyses of changes in interannual variability to decadal variability and teleconnections:
   - Model-model: past, present and future to define common model behaviors and identify the sources of model differences
   - Model-data: there are some limitations due to data interpretation, chronology, and the availability of high resolution records. But lots of work should be done in the next years. The strategy is to start from model analyses, find out the key questions and then see how the different records could help to assess the results
5. Analyses of the ocean circulation
   - Several datasets can be used. Not always ready for model assessment, but new analyses to come.
   - Test sensitivity of the THC and global climate to fresh water fluxes in the North Atlantic. This topic receives a lot of interest. It started with the CMIP/ PMIP water hosing experiment (R. Stouffer). Several groups are ready (Hadley, NCAR, CCSR (MIROC), IPSL) to run LGM hosing experiments to analyses impact of fresh water and the time scale of the response with different climate mean states. Some propositions are also emerging to have more realistic experiments such as the 8.2ka event or a deglaciation hosing experiment. This needs to be discussed because the experimental design needs to be simple and the model run < 1000 years.
6. Simulations with the carbon cycle. A proposition is emerging to have paleo-carbon experiments. It is also considered in IGBP/ AIMS

Emerging topics to address the question of " High impact or emerging uncertainties":

- Development of coupled models with interactive ice sheets
- Ice sheet and sea level: role of ice melting and estimation of sea level rise
- Floods, droughts: it would be possible to work on droughts and to assess model results by comparisons with environmental data.
- Vegetation control on emissions (eg : fires and wetlands)

Strategy:

- Data syntheses: link with IGBP/PAGES, INQUA/PALCOM
- Model data: need several workshops to address key questions, methodologies and define a set of relevant benchmarks both qualitative and quantitative (I.e how to assess the ability of climate models to reproduce changes in climate interannual to multidecadal variability)
- Modelling: key periods are already there (mid-holocene and LGM); a need is to coordinate the analyses of model results (PMIP2) + sensitivity experiments around

basic experiments.
- PMIP2 subprojects: there are about 60 projects now - synthesis of the key results is needed.

4. Improving climate models

4.1 Cloud climate feedbacks, aerosols (S. Bony)

Cloud Feedback Model Intercomparison Project – Phase 2 (CFMIP2)

The IPCC AR4 reaffirms the spread in equilibrium climate sensitivity and in transient climate response estimates among current models. Recent studies show that inter-model differences in cloud feedbacks remain the primary source of this spread, with low clouds making the largest contribution.

The main objective of CFMIP2 is to make, by the time of the AR5, an improved assessment of climate change cloud feedbacks by making progress in the

- **evaluation** of clouds simulated by climate models
- **understanding** of cloud-climate feedback processes.

An international CFMIP workshop was organized in Paris in Spring 2007 to lay the foundations for a CFMIP2 proposal, and to strengthen the links between the CFMIP, GEWEX/GCSS and the US CLIVAR Process Team communities. The CFMIP coordination committee is now composed of: M. Webb, S. Bony, G. Tselioudis and C. Bretherton.

The main activities of CFMIP2 (detailed in a full report at http://www.cfmip.net) are :

- Development of the ISCCP simulator and of a CFMIP ISCCP/CloudSat/CALIPSO simulator (CICCS) to be distributed to modelling groups to evaluate model clouds (thus contributing to the model development process) using satellite observations from the new generation of space-borne radar and lidar instruments and existing passive instruments. They are required if effective cloud-climate model metrics are to be applied to CMIP4 GCMs.
- Design and analysis of short atmosphere-only CFMIP2 experiments, requiring CICCS and other diagnostics, to better understand the physical mechanisms underlying the different cloud-climate feedbacks in climate models.
- Collaboration with GEWEX-GCSS to assess the credibility of cloud-climate feedbacks: CFMIP-GCSS CRM/LES/SCM case studies focused on the sensitivity of low-level clouds to changes in climate, process studies based on the analysis of gridpoint high frequency outputs, and development of a cloud climate metrics.

CFMIP2 wished to clarify whether WGCM continues to endorse CFMIP2 plans, and was seeking the following recommendations from WGCM (see full report at http://www.cfmip.net for details):

- Use the ISCCP simulator (and strongly encouraging the use of CICCS) in AMIP, 20C3M and 1%/yr CO_2 (plus control) experiments of CMIP4. The costs should be significantly outweighed by the scientific benefits.
- Increase the number of cloud diagnostics (currently cloud fraction is stored, but needs are for vertical structure of cloud water and ice, cloud threshold, liquid water, some of which can be derived from the ISCCP simulator) in the CMIP4 output; store 3D model outputs on model levels for a better description of irradiance (some microphysical processes that activate on certain levels are lost by interpolation), and 3D global fields daily for selected periods (eg 3-hourly to sample the diurnal cycle).
- Host the CFMIP2 experiments together with the CMIP4 archive.
- Save some high frequency instantaneous model output from selected point locations in CMIP4 and adding them to the standard output.

CFMIP2 requests ome clarification from WGCM on the plans for 1%/yr CO_2 and slab experiments in CMIP4, as this has implications for the design of the CFMIP2 experiments. Granting these requests will help to reduce systematic errors in the simulation of clouds in the present-day climate and to assess the credibility of the different cloud feedbacks produced by GCMs, making it easier to assess the reliability of climate projections.

ACTION: WGCM continues to endorse CFMIP2 plans and will recommend the use of the ISCCP-CloudSat-CALIPSO simulator (CICCS) in CMIP4 experiments. It also supports the storage of additional 3D model output, that need to be prioritised. WGCM also endorses that CFMIP2 data be hosted together with the CMIP4 archive.

4.2 Carbon cycle climate feedbacks (C. Le Quéré)

Emissions of CO_2 have been revised and updated by the Global Carbon Project. Emission estimates went up by ~0.3 GtC/yr since 1995 due to revised estimates from China. Emissions in 2006 reached 9.9 GtC/yr, well above all the IPCC scenarios generated at the end of the 1990s. For example, A1B has total emissions of 9.6 in 2006. The Global Carbon Project will be producing an updated carbon budget every year, as shown in Figure 5, from Canadell et al (2008). The distribution of emissions will be evaluated by basin. If the sources of emissions are stable, the airborne fraction will remain constant. If there is a carbon-cycle feedback, the airborne fraction can increase.

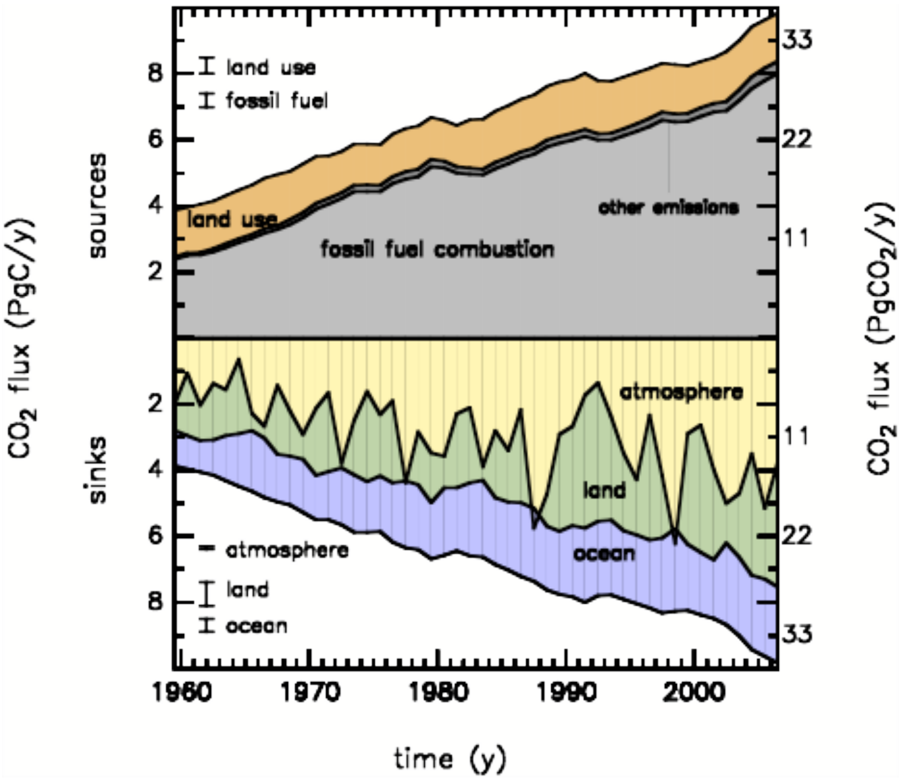


Figure 5: Global carbon budget now updated every year by the Global Carbon Project (from Canadell et al, 2008)

A positive trend in the airborne fraction of atmospheric CO_2 with p=0.90, where p is partial pressure, was detected in the observations between 1959 and 2006, but not in most C4MIP models. If this is confirmed, it suggests that the carbon-climate feedbacks may be at the upper end of the C4MIP models. Much analysis is underway to interpret the trend as a function of the rate in CO_2 emissions and of the carbon-climate feedback.

Can we represent with physical processes the trend in the observed airborne fraction of CO_2? An ocean model with geochemistry can explain the trend in the airborne fraction, with the ocean taking up progressively less CO_2 with time. This is the opposite to what happens in climate models.

A decrease in the efficiency of the Southern Ocean CO_2 sink since 1981 was detected in the observations (Le Quéré et al, 2007). This is the first estimate of a large carbon sink that is responding to climate change. Preliminary analysis suggests this decrease is underestimated in carbon-climate models. This may be due to (1) underestimated changes in Southern Ocean winds because of the absence of stratospheric ozone depletion or for other reasons, or (2) excessive vertical mixing in the Southern Ocean leading to underestimated response of the natural carbon cycle.

A global synthesis of ocean pCO_2 data is underway. It should provide information on marine trends at basin-scale since ~1970 to be used for model evaluation.

New processes being developed for carbon-climate models include: (1) N-Cycle (already working in some models), (2) land use (coming up at least in IPSL, Hadley, Frontier, GFDL), (3) fires, (4) flexible marine ecosystems, and (5) CH_4 (least advanced), (6) tropospheric ozone. A large and negative effect of tropospheric ozone on terrestrial productivity has been suggested, but projections depend heavily on future ozone scenarios. All these developments should make the models more complete and thus justify more accurate evaluation (no "missing processes" anymore).

NCAR is leading an effort to compare different surface models. Hadley and Frontier are planning ensemble simulations to address the problem of uncertainty in carbon-climate feedbacks due to model parameters. IPSL and Hadley are leading a systematic model evaluation including a comparison with atmospheric and oceanic CO_2 mean, trends, seasonal and interannual variability at various locations, terrestrial and marine productivity, and storage of passive tracers in the oceans.

In summary:

- Detection of climate-carbon feedbacks in observations is very near. Need to allow for new information to feed into the evaluation process
- Current data suggest feedbacks are on the high side
- Scenarios need to be as realistic as possible over historical period to attribute changes to emission history or feedbacks: Emission pathway over 1990-2015 is very important.

4.3 Metrics: how do we assess climate models? (K. E. Taylor)

Quantitative, scalar measures of climate model performance can be used to objectively assess the relative merits of different models, demonstrate the degree to which models are improving, and, potentially, weight individual model projections to arrive at a more reliable estimate of future climate change based on a multi-model ensemble of simulations. Plots are not considered metrics here but as diagnostic aids. Metrics can quantify errors, though are not usually targeted enough to diagnose the reasons for model errors. Active research efforts are attempting to define metrics that might be most appropriate for evaluating various aspects of model performance, in particular to quantify the fidelity of model simulations and the uncertainty in projections. The ultimate challenge will be to determine which set of performance metrics, gauging the ability of models to simulate what we observe, are particularly relevant to informing us as to the reliability of their future projections.

Most systematic, multi-variable evaluations of models focus on the highly predictable, strongly forced global pattern of the mean climate state (including annual cycle). An example is the evaluation of the climatology (1980-1999) of individual CMIP3 models, where RMS error statistics of the spatial pattern of the annual cycle were calculated for a variety of fields by summing over all grid cells and the 12 climatological months (Glecker et al, 2008). This analysis can be summarized in a single figure (Figure 6(a)), with a value of zero indicating skill equal to the mean skill of the ensemble of models. Blue (red) colors indicate a model with a smaller (larger) RMS error than the mean. In Figure 6(b), the analysis in Figure 6(a) is condensed to a single performance index (black line) and the models are ranked. The scatter of symbols (representing skill in simulating individual fields) shows that there is often little correlation between the simulation of individual fields and the performance index.

It is premature to emphasise the use of a single metric to gauge model value. The ranking will depend on which variables are included to calculate the mean performance index. A single performance index will fail to capture the complex structure of models and invites an overly simplistic interpretation, while the analysis should depend on the correct representation of the model physics. As shown by Glecker et al (2008), the skill in simulating the variance of monthly anomalies is only weakly related to the skill in simulating the climatology. The metrics will also depend on the choice of region that is analyzed, on un-forced variability in the climate system, on observational uncertainty and on the spatial resolution of the analysis (ie the coarseness of the analysis grid).

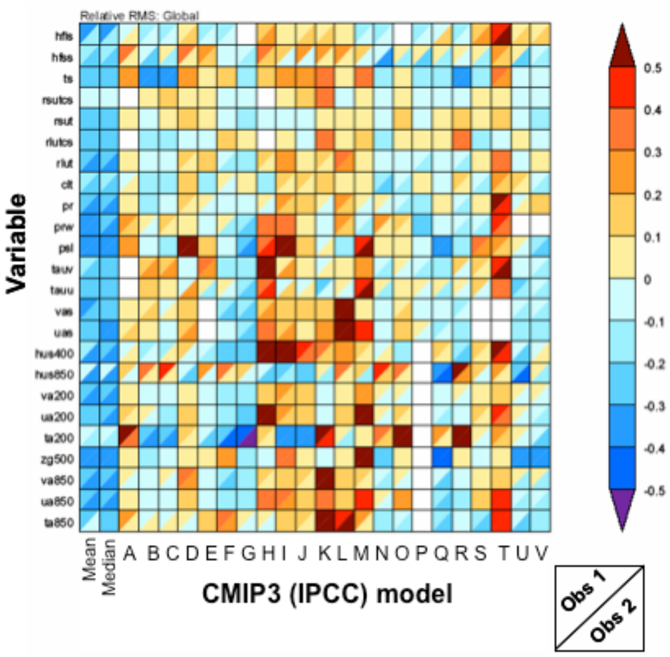
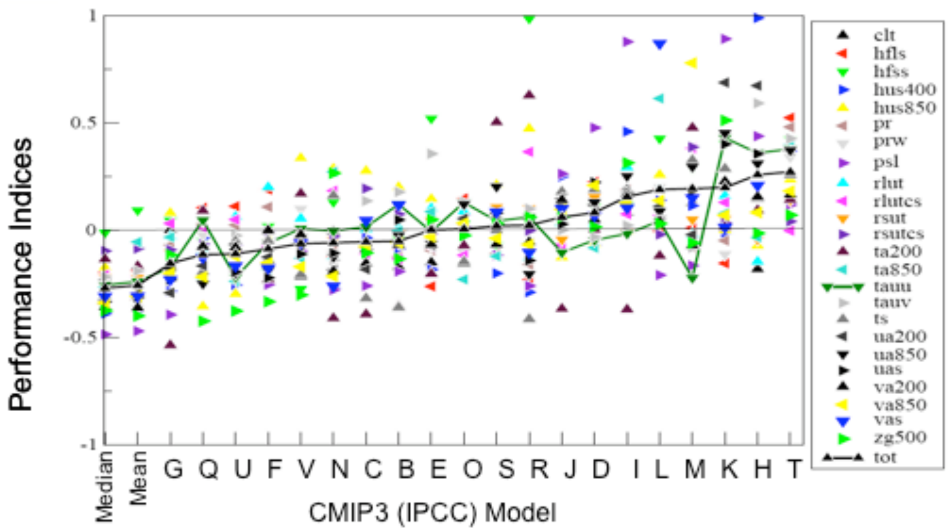


Figure 6

a) Relative error metrics for CMIP3 model simulations of the annual cycle pattern of the full global domain. Each grid square is split by a diagonal in order to show the relative error with respect to both the primary (upper left triangle) and the alternate (lower right triangle) reference datasets (from Gleckler et al, 2008).



b) Relative errors, with models ordered by the "Model Climate Performance Index," taken from Figure 4a. The indices are connected by the solid line, and the colored symbols indicate the relative error for each of the variables that contribute to the index.

The scope of metrics needs to be expanded beyond assessing the seasonal climatology and simple measures of monthly variability. We should also consider variability on other timescales (diurnal cycle, historical climate change) as well as characterizing model ability to simulate individual processes.

Future research directions for metrics research should work towards defining a comprehensive suite of performance metrics arranged in a common, hierarchical structure:

- At the lowest level: a “basket full of metrics” characterizing multiple facets of model behavior
- At a middle level: key metrics or a suite of indices characterizing various general aspects of performance (climatology, modes of variability, historical climate change)
- At the top: single “performance indices” tailored for specific applications

A minimum set of metrics should be defined that summarizes skill in simulating specific aspects of the climate system, depending on the application. Metrics should be identified in “perfect model studies”, which relate model fidelity in simulating observed phenomenon to the quality of their projections. The relationships between skill in simulating present climate and projecting the future need to be understood, bearing in mind that an accurate simulation of observables does not guarantee reliable projections. For example models that best simulate present day El Nino still have as much spread as other models in future projections. Once scientific justification for gauging confidence in projections has been established, then it will be possible to form mutli-model mean projections based on weighted individual model results.

WGCM should encourage sub-groups to develop process-orientated metrics that cover a wide range of phenomena, with PCMDI taking on a role in helping to coordinating this work and developing a hierarchy of metrics. The intent should not be to develop a ‘beauty contest’, but to enable systematic and comprehensive quantitative assessment of model performance. As understanding improves the set of metrics routinely evaluate models should evolve.

The SPARC Chemistry-Climate Model Validation Activity (CCMVal, see Section 6.3) is actively exploring the use of performance metrics in stratosphere-resolving chemistry-climate models (CCMs). CCMVal originally developed a set of process-oriented diagnostics important for stratospheric ozone as a way of assessing model performance (Eyring et al., 2005). In the CCM intercomparison that was coordinated for the 2006 WMO/UNEP Ozone Assessment by CCMVal, a set of transport metrics was explicitly used to reduce the uncertainty of the model projections of ozone recovery (Eyring et al., 2007). This was done by eliminating from the stated range of recovery dates the models that showed large differences to observed inorganic chlorine loading, since there is a direct and well understood relation between biases in maximum chlorine loading and in ozone recovery date. For the SPARC CCMVal Report that is currently being prepared in support of the 2010 WMO/UNEP Ozone Assessment, the intention is to extend the use of process-oriented metrics and to make the evaluation of the CCMs more quantitative.

5. Serving the impacts community

5.1. Regional downscaling (F. Giorgi)

The resolution of Regional Climate Models (RCMs) is currently at 15-30km, and will be increasing to 10km in the next few years, enabling the resolution of topographic and coastal effects. Most RCMs are being upgraded to a non-hydrostatic framework. Multi-decadal simulations have become the norm and some full transient simulations are available or underway.

There have been some major intercomparison projects, such as PRUDENCE, which evolved into the ENSEMBLES Project, NARCCAP, AMMA. ARCMIP and the Inter-CSE Transferability Study (ICTS). Figure 7 shows the sources of uncertainty in the simulation of temperature and precipitation change (2071-2100 minus 1961-1990) by the ensemble of PRUDENCE simulations (whole Europe) (Note: the scenario range is about half of the full IPCC range, the GCM range does not cover the full IPCC range), showing that boundary forcing is not the only source of uncertainty (Déqué et al., 2005).

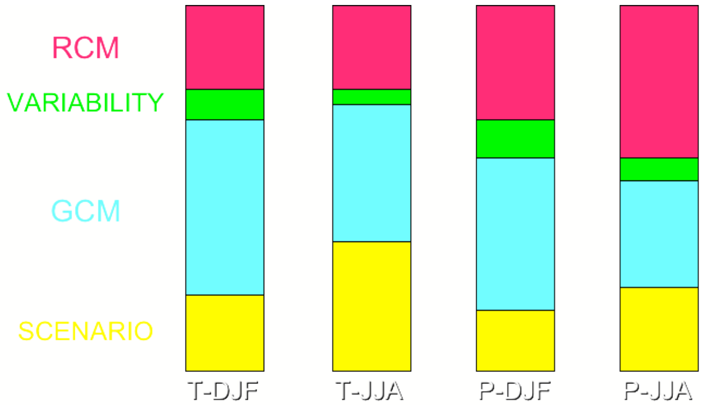


Figure 7: the sources of uncertainty in the simulation of temperature and precipitation change (2071-2100 minus 1961-1990) by the ensemble of PRUDENCE simulations (from Déqué et al., 2005)

Main technical issues currently being discussed by the RCM community are:

- Internal variability vs. external controls.
- Standard relaxation techniques vs. nudging.
- Two-way nesting – some two-way nesting systems are being developed (eg WRF), despite initial results not demonstrating much improvement.
- Development of coupled RCMs – several coupled atmosphere-ocean, atmosphere-chemistry/aerosol and atmosphere-vegetation models are under development.

There are emerging areas of applications of RCMs such as an increase in interest in their application to regional downscaling of seasonal climate forecasts and use in impact assessment studies (e.g. water resources, agriculture, health, economy, air quality, vegetation, soil erosion etc.). The use of RCMs in developing countries is being expanded by the development of portable models for different needs (The ICTP Regional Climate research NETwork(RegCM), PRECIS, RSM), enabling local scientists to help formulate a developing country perspective on climate change by conducting regional climate model experiments.

Other regionalisation techniques that are being developed including stretched grid global modelling (Stretched Grid Model Intercomparison Project (SGMIP) with four models participating (CSIRO C-CAM; Env.-Canada GEM; Meteo-France ARPEGE; NASA-GSFC GEO-3) and statistical downscaling. Progress in statistical downscaling is difficult to assess because of the large number of techniques utilizing very different methodologies often tied to local conditions. However, SD models are generally improving and some coordinated projects have been carried out or are under way (STARDEX, MICE, ENSEMBLES). There is an increasing use of statistical downscaling for impact assessment studies (AIACC, PRUDENCE, ENSEMBLES).

Both dynamical and statistical regionalisation techniques have developed enough that results are being directly applied to impact assessment studies. However, there is the need to address the lack of user knowledge in the meaning and assessment of climate projections and uncertainties in GCM fields that will be transmitted to regionalisation tools. The whole cascade of uncertainty in regional climate prediction needs to be understood. A prior assessment is needed to quantify GCM skill and the errors/uncertainty in the large scale circulation being transferred to the regional scale before using regionalisation techniques for climate change. If the errors are too great then regionalisation cannot be done.

ACTION: WCRP needs to communicate to users the need to quantify uncertainties in regionalisation (relative to other uncertainties, eg related to climate change). The peer-review process does not seem to be enough to filter out bad practice in the scientific literature. Consider putting guidelines for regionalisation on the WGCM webpage.

By its very nature, regionalisation is fragmented and heterogeneous as people are interested in different regions and applications. This means that it is difficult to set up coordinated projects, though some coordinated projects are now being implemented (Europe is at the forefront). Dynamical and statistical regionalisation techniques can be especially useful to directly engage developing country scientists into climate modelling and work needs to be done to best engage developing country scientists in climate downscaling research. Some of these points could be addressed at two ICTP-WCRP workshops being planned for November 2007 (CMIP3 results) and March 2008 (Regional Climate Modelling).

5.2. Decadal Prediction Studies (J. Murphy)

Recent studies on climate model initialisation with oceanic data include:

- Collins, 2007, Phil. Trans. R. Soc. (summary)
- Pierce et al., 2004, Clim. Change (T, S anomalies)
- Smith et al., 2007, Science (T, S anomalies)
- Troccoli and Palmer, 2007, Phil. Trans. R. Soc. (T, S absolute)
- Keenlyside et al., 2007, submitted to Nature (SST anomalies)
- Pohlmann et al., 2007, (T, S anomalies)

DePreSys (Smith et al, 2007)

A decadal climate prediction system (DePreSys) based on the Hadley Centre coupled global climate model, HadCM3, has been developed at the Met Office Hadley Centre (Smith et al, 2007). On decadal time scales, climate could be dominated by internal variability arising from unforced natural changes in the climate system such as El Niño, fluctuations in the thermohaline circulation, and anomalies of ocean heat content. This could lead to short-term changes, especially regionally, that are quite different from the mean warming expected over the next century in response to anthropogenic forcing. The model is forced by greenhouse gases and sulphate aerosols (SRES B2 scenario – intermediate changes).

Accurate initialisation of the state of the ocean is therefore important for a decadal climate prediction system using coupled climate models. Climate drift during forecasts is avoided by assimilating observed ocean temperature and salinity anomalies, rather than observed values, to the model climate. The atmospheric component of HadCM3 is initialised by relaxing (with a 3 hour timescale) the horizontal winds, potential temperature and surface pressure to the ERA-15 reanalysis.

Hindcasts were started from the 1st March, June, September and December in each year from 1982 to 2001, each 10 years long with 4 ensemble members started on consecutive days. The impact of initial conditions is assessed the root mean square error of the ensemble mean of DePreSys to a run with no assimilation (NoAssim) forced by the same external forcing though without the assimilation of observed initial conditions. Improved skill is found in hindcasts of global mean surface temperature (Ts), shown in Figure 8(a), and of upper ocean heat content (H), shown in Figure 8(b). Linear regression coefficients that relate the state of El Niño, as measured by SST in the Niño3 region to Ts are computed from the transient HadCM3 simulations. Improved skill of Ts can be explained mainly by the skill in predicting ENSO in first 15 to 18 months, but not at longer lead times (dashed line, Figure 8(a)). The DePreSys hindcast has a warm bias in Ts compared to the NoAssim hindcast that has been attributed to the interdecadal variability of the upper ocean heat content (Smith et al, 2007). Removing this bias leads to reduced skill so that the RMSE of the DePreSys hindcast is not significantly different to that of NoAssim (dotted curve, Figure 8(a)). The increased predictive skill of DePreSys longer lead times therefore results mainly from the initialisation of the low-frequency upper ocean heat content variability. Improved skill regionally is also attributed to the initalization of the ocean heat content that leads to improved coupled feedbacks that are absent in the NoAssim hindcasts (Smith et al, 2007).

A decadal forecast with DePreSys was initiated from June 2005. Internal variability in the DePreSys offsets the effects of anthropogenic forcing in the first few years, leading to no net warming before 2008. In contrast, the equivalent NoAssim forecast warms during this period (see Smith et al, 2007 for more details).

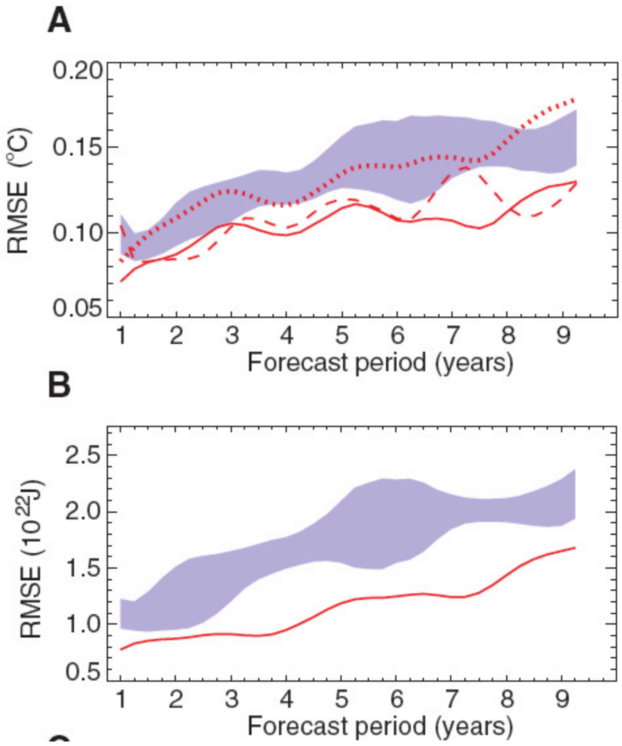


Figure 8 (from Smith et al, 2007).
(A): RMSE of globally averaged annual mean surface temperature (Ts) anomalies (relative to 1979–2001) as a function of forecast period. DePreSys (solid red curve) is compared with the NoAssim hindcasts (the blue shading shows the 5 to 95% CI region where differences between DePreSys and NoAssim are not significant). The dashed red curve shows the effect of removing from the DePreSys hindcasts differences between DePreSys and NoAssim that are linearly attributable to the state of El Niño. The dotted red curve shows the effect of removing from the DePreSys hindcasts the mean difference between DePreSys and NoAssim hindcasts of Ts for the coming 9 years. Observations are taken from the HadCRUT2vOA dataset.
(B): As (A), but for ocean heat content in the upper 113m (relative to 1941–1996). Observations of ocean heat content are computed from analyses of ocean temperature observations.

Further analysis is required in order to assess the predictability of other climate variables, including precipitation and extreme events, and to present results in a probabilistic framework. The gap between the actual skill and the theoretical skill diagnosed from the intra-ensemble correlation suggests that improving the model and its initialisation would give more accurate forecasts. Efforts to improve the model are ongoing and methods for improving initialisation, by achieving more balanced initial conditions and including additional observations (such as altimeter data), will be explored. The ensemble technique used so far accounts for the influence of uncertainties in the initial conditions but not for modelling uncertainties. The possibility of generating ensembles which account for modelling uncertainties, through perturbations to the model physics, will be investigated.

Recommendations:

- Intitialise the ocean component with anomalies, though this requires a long term transient simulation to determine the model climatology. It is unrealistic to start short term runs from pure data as drift in the model climate will dominate the results.
- Initialising from just SST gives inconsistent results with regards to salinity. Both are needed to avoid significant drift, though Keenlyside et al (2007) found that there is sufficient skill initializing from SST only. The current ARGO era means that initializing with full depth T-S goes far to avoid drift.

6. High impact or emerging uncertainties

6.1. Ice sheets and sea level (J. Gregory)

The largest uncertainty in projections of sea level rise in the IPCC WG1 AR4 relates to changes of mass of the ice sheets of Greenland and Antarctica. Ice sheet mass balance is the sum of surface mass balance (mostly solid precipitation and melting leading to runoff) and ice discharge into ice shelves and icebergs. Surface mass balance depends on surface climate. Ice discharge depends on ice flow, which occurs through deformation of ice and sliding at the bed. Ice flow is slow over the majority of the ice sheet area (a few metres per year), such that the renewal time for the ice sheet (volume divided by surface mass balance) is many millennia. Models exist for "slow" dynamics. They are able to account for the present ice sheet topography and the long-term evolution of the ice sheets during glacial cycles.

Although these models can reproduce the ice sheet flow and discharge averaged over large areas, they do not simulate the detail of ice flow on the km scale. The majority of discharge actually occurs through fast flowing ice streams and outlet glaciers, which occupy a small fraction of the ice sheet area, where the velocity can be up to km per year. The flow in the ice streams is rapid because of sliding, due to readily deformable sediment and lubrication by melt water at the bed. Accelerated flow has been observed in recent years in some Amundsea Sea ice streams and many Greenland outlet glaciers, leading to increased discharge and positive SLR contributions from both ice sheets, assessed by the AR4 as 0.4+-0.4 mm yr-1 in total during 1993-2003.

This is a small fraction of the current rate of sea level rise of ~3 mm yr-1, but the observed acceleration gives rise to large uncertainties in projections because there is presently only limited understanding of the controls on ice-stream flow. Weakening or removal of ice shelves because of ocean warming and access of
increased surface melt water to the ice sheet bed may be the causes of recent acceleration. We currently do not have either an empirical or a model method of making projections of these effects. As an illustration, if the 1993-2003 contribution were to scale up linearly with global average temperature change, it would add up to 0.2 m to AR4 projections for the 21st century cf A1B projection of 0.21-0.48 m. Some think this is an implausibly large contribution, others think that the effect could be far larger! There is no consensus, especially beyond the 21st century.

In order to address the problem, we need:

- Models of ice-stream flow to be developed and included in continental ice-sheet models. Such development has to consider grounding-line migration, the buttressing effect of ice-shelves, and the lubrication of the bed by surface melt water.
- Ocean models to include the interaction with ice, leading to melting and freezing at the grounding line and the underside of ice shelves. This may entail rather high resolution for simulating ocean circulation on the continental shelves and in sub-ice cavities and involves including a solid upper boundary in the ocean model.
- Models of ice-shelf surface mass balance and the way in which surface melting may cause disintegration of ice shelves.
- There is work on all of these areas in the glaciological community, but such models are not in an advanced state, and have not been coupled to GCMs.

Although ice-sheet dynamics is the most important uncertainty, there are several other important problems with sea level projection:

- Ice sheet surface mass balance is not well simulated by AGCMs, which do not have sufficient resolution or surface schemes for ice. The large uncertainty (range of 1.6-4.6degC) for the threshold of viability of the Greenland ice sheet results from surface mass balance uncertainty, including the relation of regional ice sheet climate change to global climate change. In recent years Greenland has warmed more rapidly than AOGCM simulations, while Antarctica has not warmed (except on the Peninsula), in disagreement with simulations. This is an important issue because it means the recent observed probable *loss* of mass by the Antarctic ice sheet has the opposite sign to simulations, which suggest that it should *gain* mass during the 21st century because of greater snowfall in association with warming. Possibly local ozone-induced or other changes may be affecting the regional climate. Will they continue to do in coming decades?

- Glaciers and ice caps are expected to give the second-largest contribution to sea level rise in this century (after thermal expansion). They are much too small to be simulated directly in GCMs, so aggregate treatments are needed. Existing models have substantial uncertainties, because only a few dozen of the >100,000 G&IC have been modelled in detail. A global synthesis of observational datasets was done for the first time in the AR4, and suggests that existing models underestimate the present rate of mass loss. In coming decades a first-order effect will be the loss of area caused by glacier retreat, but only simple models of this are available.
- Improvements in observational estimates since the TAR have led to a more certain assessment of the sea level budget in the AR4. This has shown more definitely that there is a gap in the budget; the sum of known terms (thermal expansion, G&IC and ice sheets) is less than the observed rate of sea level rise. For 1961-2003 the observed rate of SLR is 1.8+-0.5 mm yr-1, of which 1.1+-0.5 mm yr-1 is explained and 0.7+-0.7 mm yr-1 is not. The largest terms are expansion and G&IC, and for these models and observations agree. Hence this is not principally a model problem; it is a problem with understanding. The possibilities are that the observed rate of SLR is overestimated, there is a term we haven't thought of, or one of the known terms is underestimated in both models and observations. At present there isn't support for any of these in the literature. The existence of a gap implies the SLR projections may be an underestimate. If the gap is a real addition but not climate-related and continued unchanged, it would add 0.07 m to SLR during the 21st century, which is not a lot. On the other hand, if it scaled up with climate change, it would add a great deal.

In response to the question of how the climate modelling community move forward on the issue of the role of ice sheets in climate change, J. Gregory responded that climate centres should include ice sheet modelling, for example using GLIMMER (http://forge.nesc.ac.uk/projects/glimmer), a community three-dimensional thermomechanical ice sheet model, designed to be interfaced to a range of global climate models. GCMs need more glacial expertise, though it has not necessarily been proved yet that there is benefit in coupling with GCMs, rather than running off line simulations.

ACTION: Maintain links to ice sheet modelling community through J. Gregory, as well as maintaining interactions with CliC.

6.2. Extreme events

The CLIVAR Expert Team on Climate Change Detection and Indices (ETCCDI) is in the process of revisiting the list of indices of extremes calculated for AR4, and suggest updates/ improvements based on experience now, in time for implementation for AR5.

6.3. Air quality and climate change: SPARC Perspective on Chemistry, Air Quality and Climate (V. Eyring)

Chemistry is important in AGCMs and ESMs for a number of reasons. Much of human induced climate forcing occurs through chemically active species (N_2O, CH_4, halogens), whose atmospheric lifetime is controlled by transport and chemistry. Chemistry also affects the radiative budget through its impact on the formation of aerosols. Ozone is a naturally occurring greenhouse gas with a relatively short lifetime, and hence is highly variable in both space and time, with its distribution controlled by transport and chemistry. In particular there is a strong contrast in ozone abundance across the tropopause, which has important climate effects both direct (through radiation) and indirect (through dynamics). Stratospheric ozone has been subject to a major perturbation since the late 1970s due to anthropogenic emissions of ozone-depleting substances, now controlled under the Montreal Protocol, and it is necessary to account for the climate effects of ozone depletion and recovery in order to correctly detect and attribute greenhouse-gas induced climate change. The stratospheric ozone layer also provides the primary mechanism for the effects of solar variability on climate. Finally, chemistry is the link between climate and air quality, and a major impact of climate change that needs to be quantified involves changes in air quality (and air quality standards).. The following are the main important chemistry-climate interactions that are currently being addressed within WCRP:

Effect of stratospheric ozone recovery on tropospheric climate
This is being addressed by by SPARC CCMVal (http://www.pa.op.dlr.de/CCMVal). Over the past 30 years, ozone has been depleted while climate change has progressed, though recovery is predicted. Over the next half-century, ozone recovery is predicted. The projected stratospheric ozone evolution in the 21st century on a global scale is mainly determined by decreases in halogen concentrations and continued cooling of the global stratosphere due to increases in greenhouse gases. Over sub-global regions, ozone is also affected by stratospheric circulation changes arising from climate change. For example, models consistently project a decrease in tropical lower stratospheric ozone associated with increased tropical upwelling (Eyring et al., 2007). Such a decrease in lower stratospheric tropical ozone is in fact observed (Randel and Wu, 2007), but it should probably be attributed to climate change, not to CFCs, and so is not expected to reverse in the future.

In addition to the direct radiative effects of stratospheric ozone on the energy balance of the atmosphere (which mainly come from ozone in the lower stratosphere), ozone changes can induce stratospheric circulation changes in the lower stratosphere which can then affect tropospheric weather and climate, especially at high latitudes. Such an effect is manifested in the surface response to the ozone hole, where observations and model results reveal falling geopotential heights poleward of 60°S and rising geopotential heights in the middle latitudes during summer months (following the breakdown of the stratospheric polar vortex). This is associated with a trend in the Southern Annular Mode during the summer months that is accompanied by significant cooling over most of Antarctica (Gillett and Thompson, 2003). A recent study with the GEOS-CCM finds that this trend is projected to reverse in the future during the summer months as the ozone hole recovers. In contrast, the GCMs used for the IPCC AR4 project no reversal of the trend. Those GCMs which included some representation of ozone recovery show a weakening of the trend, but no reversal (Perlwitz et al. 2008). This suggests that stratospheric processes, including chemistry, are crucial for a reliable estimate of tropospheric changes in the future (especially at high Southern latitudes), and should be included in future AGCM and ESM integrations. In addition to making the ozone and dynamical fields self-consistent, interactive modeling of stratospheric ozone would also alleviate the current difficulties of merging independent characterizations of tropospheric and stratospheric ozone from different models and forcing scenarios.

Tropospheric chemistry and climate
A changing climate will change air quality and the tropospheric ozone budget has a role in climate change. These questions are being addressed by the Atmospheric Chemistry and Climate (AC&C) joint initiative of IGBP-IGAC and WCRP-SPARC in the Activities 1 (coordinated by P. Hess) and 4 (coordinated by D. Shindell and J. F. Lamarque) – http://www.igac.noaa.gov/ACandC.php. The tropospheric ozone burden has increased by 71 Tg between 1890 and 1990 — an increase of ~30%. In future climates the decreased tropospheric burden is the result of a competition between increased ozone destruction due to higher relative humidity and increased influx of ozone from the stratosphere. Stevenson et al. (2006) show that different models have different sensitivities to these processes. In polluted regions, climate change will have a positive feedback on surface ozone, whereas in clean regions, climate change will have a negative feedback on surface ozone.

Aerosols in a changing climate
This question is being addressed by the Global Aerosol Model Intercomparison Project (AeroCom - http://nansen.ipsl.jussieu.fr/AEROCOM/). Differences in aerosol mass depend largely on differences of model-specific transports, parameterizations of aerosol interactions, microphysical processes, and to a lesser extent on their (precursor) emissions. The questions being addressed are how well aerosol transport between any source and receptor can be quantified, and how we can improve the evaluation of aerosol process parameterizations.

SPARC CCMVal and AC&C contributions to coordinated AOGCM and ESM experiments
Since it is unrealistic to imagine that all coupled models will have interactive chemistry by 2010, CCMVal could provide best guess plus uncertainties in the stratospheric ozone distribution for coupled models that do not have the capacity to include interactive stratospheric ozone chemistry. For tropospheric chemistry, AC&C could provide emissions for non CO_2 gases. The proposed list of species is: CH_4, N_2O, CFCs, HCFCs, SF6, CO, NO_x, VOCs, SO_2, OC, BC, NH_3, H_2 (AC&C Activity 4, in collaboration with the IPCC community on future emissions). Emissions should be provided by source and region for air quality studies.

Whether interactive chemistry needs to be run in climate models to study the impact of climate change on air quality depends on the region. For example, impacts on the polar vortex lead to coupled feedbacks that need to be resolved. CCMVal and AC&C could provide metrics and an evaluation standard for the representation of chemistry in the coupled models.

Requirements/Needs:
- Regional emissions are needed because of the nonlinearities in chemistry and strong regional impacts of aerosols
- Climate impact of forcings depends on its localization
- VOC speciation is important because of the different rates of ozone production and aerosol formation

Challenges:
- Ability to simulate air quality requires the most detailed emissions (short-term simulations) and chemical mechanisms
- Aerosol modelling requires knowledge of precursor emissions and chemistry at regional scale
- Representation of aerosol indirect effects
- Representation of chemical impacts on vegetation (ozone, nitrogen deposition)
- Importance of landuse/landcover for biogenic and biomass burning emissions
- Consistency between present-day and future emissions
- Implementation of detailed sub-models, or components, of the stratosphere, atmospheric chemistry, and the aerosol cycle in AOGCMs and ESMs.

7. Meeting summary

Collaboration with other WCRP Projects and WG II and III
- GEWEX – link through S. Bony and G. Tselioudis. Should WGCM be more involved in land processes? GEWEX can prepare a list of cloud metrics for model performance evaluation to submit to WGCM at the 4th Pan-GCSS Meeting: Advances on Modelling and Observing Clouds and Convection to be held in Toulouse in June 2008.

ACTION: GEWEX to prepare list of cloud metrics for WGCM (S. Bony)

- SPARC – what is being done with regards to ozone, other than what CCCma is planning in terms of a 3D ozone dataset?

- WGSIP – Testing models in SP mode, for example, based on CHFP Experiment protocol that has been designed as a protocol to test the state-of-the-art of SP, against which climate models can be tested. Which models will be put through the CHFP, or would the community prefer a lighter weight version of the experiment? M. Kimoto and A. Hirst (contact O. Alves) have in indicated that their groups would be interested. On the other hand, J. Mitchell said that the Met Office would probably not participate as it is using the same model in SP and climate mode.

ACTION: Inform WGCM of Climate-system Historical Forecast Project (CHFP) protocol once webpage is up to date (A. Pirani)

ACTION: Gather response from modelling groups on interest in participating in Climate-system Historical Forecast Project (CHFP) once webpage is ready to be distributed.

- Global Carbon Project – can make recommendations on 1% historical land use.

ACTION: Contact J. Fedema to bring historical land use up at Amsterdam scenarios meeting and will feedback to GCP if useful (G. Meehl).

Emissions Scenarios and Coordinated Experiments

ACTION: Place scenario strategy onto WGCM webpage (G. Meehl to A. Pirani).

How will the Scenarios community make the transition from the 20th Century to the 21st Century?

Which chemical species are needed to drive models in prescribed concentrations mode and emissions mode?

ACTION: See list of chemical species prepared from modelling centre consultation. Note that ozone is not on the list. CCCma has stated that they will supply ozone dataset. Give list to K. E. Taylor to take to Amsterdam meeting (G. Meehl).

ACTION: Link to atmospheric chemistry activities by passing chemical species list, response from modelling centres and CCCma O_3 plans to AC&C coordinators (P. Rasch, S. Doherty and A. R. Ravishankara) (G. Meehl).

Improving Climate Models

- Cloud Feedbacks – WGCM endorses CFMIP2, together with supporting the archiving of CFMIP2 data with the CMIP4 archive. Plans for a 1% /yr and slab experiments are also endorsed.

  Cloud simulator (CICCS) will be delivered by the end of 2008 to be used by the community in CMIP4. It will be compatible with the current ISCCP simulator and will be part of a common framework supporting all simulators. WGCM agrees in principle with the proposed increase in cloud diagnostics, with higher frequency output for some regions and locations, and for 3D model output to be stored on model levels.

ACTION: Prepare a list of cloud diagnostics for the next WGCM meeting (S.Bony).

G. Tselioudis (representing GEWEX) strongly supports maintaining a GCSS-CFMIP collaboration with S. Bony attending GCSS meetings and G. Tselioudis attending WGCM meetings.

- Carbon Cycle Feedbacks

ACTION: After further iterations (C. Le Quéré, K. E. Taylor, J. Gregory) give details on final outcome of experimental design to assess carbon cycle feedbacks (to A. Pirani).

- Metrics

ACTION: PCMDI to develop a metrics webpage (K. E. Taylor).

High Impact and Emerging Uncertainties

- Ice Sheets – Encourage climate modelling groups to work with ice modelling groups. Develop the link between WGCM and the ice sheet modelling community, for example through J. Gregory. WGCM also needs to link to the PMIP and CliC communities.

ACTION: Put list of ice modelling activities from J. Gregory's presentation onto WGCM webpage (A. Pirani).

- Air Quality – WGCM needs to maintain its link to SPARC. A recommendation should be make in terms of which scenario should be used for SPARC atmospheric simulations that were originally based on the A1.2 (intermediate) scenario.

ACTION: Communicate with SPARC community at the SPARC SSG held in Bremen, Germany on 18 to 21 September 2007 (M. Giorgetta).

Need to communicate decision to leave inclusion of air quality as an option for short term experiments, without making this a standard component due to modelling groups having different chemistry capabilities.

ACTION: Follow up with AC&C with regards to the inclusion of air quality in the short term experiments (G. Meehl).

Serving the impacts community

- Regional downscaling – WGCM is in the best position to provide recommendations to the user community, including the limitations of downscaling such as the fact that the current range of climate models may not be enough to quantify uncertainty. The T. Carter recommendations should be re-evaluated. The regional interest in climate change is the ultimate driver for government action. WGCM needs to make a statement of the current strength, weaknesses, progress and need for regional downscaling of climate projections, and agree it with TGICA.

ACTION: Update recommendations for regional climate modelling at WCRP Modelling Summit (F. Giorgi).

ACTION: Prepare a WGCM statement on the state-of-the-art of regional climate modelling (J. Mitchell).

ACTION: Develop a WGCM webpage gathering information on regional climate modelling (A. Pirani).

ACTION: Need to increase interactions with the wide community of stakeholders and end users of CMIP3 climate projections, to reduce likelihood of misinterpretation of model output and to identify specific needs that might be addressed in future CMIP phases. Initially consult with T. Busalacchi, C. Rosenflag (agricultural community) and M. Cane who have experience in this.

- Decadal prediction – Proposal being developed by T. Stockdale and G. Hegerl has been commissioned by WGCM. Further iterations are needed and a subgroup has been formed between WGCM and WGSIP. T. Stockdale suggests that this proposal replace the proposal presented to the JSC by T. Palmer due to its wider scope.

ACTION: WCGM proposes members for a WGCM-WGSIP decadal prediction subgroup (led by G. Hegerl and T. Stockdale, J. Murphy, G. Hergel, R. Stouffer, G. Meehl, M. Kimoto M. Giorgetta) to prepare a revised proposal which should be sent to this group as well as WGCM in general. The JSC group assigned for decadal prediction also needs to be informed.

ACTION: Explore the possibility of WGCM and WGSIP jointly holding a small decadal workshop.

WGCM Membership

WGCM recommends that the terms of those due for membership renewal in 2008 be extended. WGCM is looking for membership from the atmospheric chemistry community and recommends V. Eyring. GFDL will continue to be represented by R. Stouffer.

Next WGCM Meeting
Tentative dates for the next meeting are the 3rd or 4th weeks of September or late October. There is the possibility of holding the next meeting joint with the next IGBP/AIMES meeting.

ACTION: Explore the possibility of holding a joint WGCM-AIMES meeting (G. Meehl).

References

Collins, M., 2007, Ensembles and probabilities: a new era in the prediction of climate change. *Philosophical Transactions of the Royal Society A*, 365 (1857), 1957-1970.

Déqué, M., R.G. Jones, M. Wild, F. Giorgi, J.H. Christensen, D.C. Hassell, P.L. Vidale, B. Rockel, D. Jacob, E. Kjellström, M. de Castro, F. Kucharski and B. van den Hurk, 2005: Global high resolution versus Limited Area Model climate change projections over Europe: quantifying confidence level from PRUDENCE results. *Clim. Dyn*., **25(6)**, 653-670.

Eyring V., N.R.P. Harris, M. Rex, T.G. Shepherd, D.W. Fahey, G.T. Amanatidis, J. Austin, M.P. Chipperfield, M. Dameris, P.M. De F. Forster, A. Gettelman, H.F. Graf, T. Nagashima, P.A. Newman, S. Pawson, M.J. Prather, J.A. Pyle, R.J. Salawitch, B.D. Santer, and D.W. Waugh, 2005: A strategy for process-oriented validation of coupled chemistry-climate models, *Bull. Am. Meteorol. Soc.*, **86**, 1117–1133.

Eyring, V., D. W. Waugh, G. E. Bodeker, et al., 2007, Multimodel projections of stratospheric ozone in the 21st century, *J. Geophys. Res*., **112**, D16303, doi:10.1029/2006JD008332.

Hibbard, K. A. , G. H. Meehl, P. Cox, and P. Friedlingstein, 2007: A strategy for climate change stabilisation experiments, *EOS*, **88**, 217-221.

Gillett, N.P. and D.W.J. Thompson, 2003: Simulation of Recent Southern Hemisphere Climate Change, *Science*, **302**, 273-275

Gleckler, P.J., K.E. Taylor, and C. Doutriaux, 2008: Performance metrics for climate models. *J. Geophys. Res*., **113**, D06104, doi: 10.1029/2007JD008972.

Griffies, S. M., A. Biastoch, C. Böning, F. O. Bryan, G. Danabasoglu, E. P. Chassignet, M. H. England, R. Gerdes, H. Haak, R. W. Halberg, W. Hazeleger, J. Jungclaus, W. G. Large, G. Madec, A. Pirani, B. L. Samuels, M. Sheinert, A. Sen Gupta, C. A. Severijns, H. L. Simmons, A. M. Treguier, M. Winton, S. Yeager, and J. Yin, 2008: A Proposal for Coordinated Ocean-ice Reference Experiments (COREs). *Ocean Modelling*, submitted.

Josep G. Canadell, Corinne Le Quere, Michael R. Raupach, Christopher B. Field, Erik T. Buitenhuis, Philippe Ciais, Thomas J. Conway, Nathan P. Gillett, R. A. Houghton, and Gregg Marland, 2007: Carbon sink slowdown contributing to rapid growth in atmospheric CO2, *Proceedings of the National Academy of Sciences*, 10.1073/pnas.0702737104, (Online Early Edition).

Jung, T., S. K. Gulev, I. Rudeva, and V. Soloviov, 2006: Sensitivity of extratropical cyclone characteristics to horizontal resolution in the ECMWF model. *Q. J. Royal Meteorol. Soc.*, **132**, 1839-1857.

Keenlyside, N.S., M. Latif, J. Jungclaus, L. Kornblueh, and E. Roeckner, 2007: Forecasting North Atlantic Sector Decadal Climate Variability. *Nature*, submitted.

Le Quéré et al, 2007: Saturation of the Southern Ocean CO2 Sink Due to Recent Climate Change, *Science*, **316** (5832),1735-1738, DOI:10.1126/science.1136188

Meehl, G.A., C. Covey, T. Delworth, M. Latif, B. McAvaney, J.F.B. Mitchell, R.J. Stouffer, and K.E. Taylor, 2007: THE WCRP CMIP3 Multimodel Dataset: A New Era in Climate Change Research. *Bull. Amer. Meteor. Soc*., **88**, 1383-1394.

Meehl G. A., and K. A. Hibbard, 2007: A Strategy for Climate Change Stabilisation Experiments with AOGCMs and ESMs, WCRP Informal Report N. 3, ICPO Publication N. 122, IGBP Report N. 57.

Otto-Bliesner et al. 2006: Simulating Arctic Climate Warmth and Icefield Retreat in the Last Interglaciation, *Science*, **311** (5768), 1751-1753, DOI: 10.1126/science.1120808.

Perlwitz J., S. Pawson, R. L. Fogt, J. E. Nielsen, and W. D. Neff, 2008: Impact of stratospheric ozone hole recovery on Antarctic climate, *Geophys. Res. Lett.,* **35**, L08714, doi:10.1029/2008GL033317.

Pierce, D. W., Barnett, T. P., Tokmakian, R., Semtner, A., Maltrud, M., Lysne, J., and Craig, A., 2004: The ACPI Project, Element 1: Initializing a Coupled Climate Model from Observed Conditions, *Clim. Change* **62**, 13–28

Pohlmann, H., Jungclaus, J., and Marotzke, J., 2007: Initialization of a global climate model with oceanic reanalysis, in preparation.

Randel W. J. and F. Wu, 2007: A stratospheric ozone profile data set for 1979–2005: Variability, trends, and comparisons with column ozone data, *J. Geophys. Res.*, **112**, D06313, doi:10.1029/2006JD007339.

Smith, D. M., S. Cusack, A. W. Colman, C. K. Folland, G. R. Harris, J. M. Murphy, 2007: Improved Surface Temperature Prediction for the Coming Decade from a Global Climate Model, *Science*, **10**, 796-799, DOI: 10.1126/science.1139540

Stevenson, D.S., *et al.*, 2006: Multi-model ensemble simulations of present-day and near-future tropospheric ozone, *J. Geophys. Res*., **111**, D08301, doi:10.1029/2005JD006338.

Troccoli, A. and Palmer, T. N., 2007: Ensemble decadal prediction from analysed initial conditions. *Phil. Trans. R. Soc. A,* **365**, doi:10.1098/rsta.2007.2079.

Wu Q, and Karoly D.J., 2007, Implications of changes in the atmospheric circulation on the detection of regional surface air temperature trends. *Geophysical Research Letters*, **34**(8): L08703.

APPENDIX I

WGCM-11 Agenda
(Hamburg, Germany, 3-5 September 2007)

DAY 1 - Monday, September 3

0900-0910 Welcome - J. Mitchell, G. Meehl
-Introductions
-Times, local arrangements - M. Giorgetta
-Explanation of Agenda - J. Mitchell
Adoption of Agenda

0910-0930 Reports and news from governing groups (JSC/CLIVAR) (5 minutes each)
-JSC-XXVIII session, Zanzibar, Tanzania - G.Flato
-CLIVAR SSG session and International CLIVAR Project Office - G. Meehl
-ACC: reports from Beijing, Paris Meetings - G.Flato or other JSC Member/J. Mitchell
-Modelling Summit - TBD

0930-1030 Updates on proposed coordinated experiments and science input relevant to the science questions related to the short term and long term experiments (follow-up to WGCM Victoria meeting, 2006) (15 minute talks, 5 minutes questions)
-Review proposed strategy for coordinated experiments addressing short term and longer term climate stabilization (note ENSEMBLES activity) - G. Meehl
-Review responses from modelling groups and input from others in the community (e.g. IDAG) regarding experimental design - G.Meehl, G. Hegerl
-Other possible experiments to test feedbacks in the experimental design - K. Taylor, P. Friedlingstein

1030-1050 Coffee break

-WGSIP input on the short-term decadal experiments, issues related to coupled initialization, ENSO in AOGCMs, and other issues related to WGCM - T. Stockdale
-Perspective on the long term stabilization experiments, status of benchmark stabilization concentration scenarios (Snowmass meeting outcomes, preview upcoming Netherlands meeting, including relevant input on proposed experimental design from TGICA and scenarios consortium, interface with WG2 community) - N. Nakicenovic
-Working Group on Numerical Experimentation WGNE perspective on coordinated experiments, especially the higher resolution short-term experiments, and other issues directly related to WGCM - M.Miller/K. Taylor
-Working Group on Ocean Modelling (WGOMD) including any input on proposed short term and long term coordinated experiments, appropriate resolution, etc. - H. Banks
-Coupled initialization and associated science questions related to the short term decadal prediction problem - J. Murphy

1230-1345 Lunch

1345-1530 Discussion of issues for WGCM (leader in brackets) (15 minute talk, 5 minute questions)
Improving climate models
-Cloud climate feedbacks, aerosols - (S.Bony)
-Carbon cycle climate feedbacks:
(1) long term stabilization experiments
(2) strategy for reduction of uncertainties - (C. Le Quéré)

-Metrics: How do we show how well (or poorly) climate models work - K.Taylor

High impact or emerging uncertainties

-Ice sheets and sea level – (J. Gregory)

-Air quality and climate change - (V. Eyring)

1530-1600 Coffee break

1600-1720 Discussion of issues for WGCM (leader in brackets) (15 minute talk, 5 minutes questions, cont'd)

Serving the impacts community:

-Regional downscaling for short term and long term (F. Giorgi)

-Decadal prediction - latest developments in the modeling community, and application to impacts - (M. Giorgetta)

-Impacts and the next set of coordinated experiments (lead?)

-WGCM's input to WCRP SBSTA paper on Climate Modelling needs for regionalization - (A.Henderson-Sellers)

1720-1745 Recap of day's session and discussion of topics to be re-visited in the next 2 days

DAY 2 - Tuesday, September 4

0900-1030 Updates from modelling centres (including where are the modeling groups in their development cycles related to the proposed experimental design, both classes of models or just one, computing issues related to the large number of experiments proposed) (12 minutes each, 3 minutes questions)

-Germany - M. Giorgetta

-France - P. Bracannot/S. Bony

-Canada - G. Flato

-Italy - F. Giorgi

-Australia - A. Hirst

-Japan - M. Kimoto

1030-1100 Coffee Break

-U.K. - J. Mitchell

-NCAR - G. Meehl

-GFDL - S. Griffies

1145-1230 WGCM activities

Coupled Model Intercomparison Project (CMIP) - R.Stouffer, C.Covey, K.Taylor, D. Bader

-Current status of CMIP3 multi-model dataset and analysis efforts

-CFMIP next steps (anything not discussed under clouds and aerosols)

-How will model output from the next set of coordinated experiments be archived and accessed?

1230-1345 Lunch break

1345-1405 WGCM activities (continued)

Paleoclimate Modelling(including relative to climate sensitivity, sea level if not already discussed) - P. Braconnot

1405-1545 Discussion topics

Role of EMICs to complement AOGCMs and ESMs, especially to interpolate between benchmark scenarios; pattern scaling – (J. Mitchell lead)

Computer resources required for coordinated experiments - (G. Meehl lead)

1545-1615 Coffee break

1615-1715 Discussion topics (cont.)
Revisit experimental design in light of outstanding science or logistical considerations - (J. Mitchell lead)

1715-1745 Recap of day's session, and discuss topics for discussion Wednesday

Evening Dinner hosted by MPI

DAY 3 - Wednesday, September 5

0900-1030 Review of decisions, requests to projects etc, and future directions

1030-1100 Coffee break

1100-1230 Review linkages with WGSIP, WGNE, AIMES, WG2, WG3, and others

1230-1345 Lunch break

1345-1530 Recap of session, and re-visit any topics that need further consideration

1530-1600 Coffee break

1600-1700 Closed session
-Membership issues
-Next Session: venue, dates

1700 End of WGCM-11 Session

APPENDIX II

Participants at WGCM-11
(Hamburg, Germany, 3-5 September 2007)

WGCM Co-chairs

Professor J. MITCHELL
Met Office
Fitzroy Road
Exeter EX1 3PB
UK
Tel: +44-1392-884604
Fax: +44-1392-884400
Email: john.f.mitchell@metoffice.gov.uk

Dr G.A. MEEHL
National Center for Atmospheric Research (NCAR)
Climate and Global Dynamics Division
P.O. Box 3000
Boulder, CO 80307-3000
USA
Tel: +1-303-497-1331
Fax: +1-303-497-1333
Email: meehl@ncar.ucar.edu

WGCM Members

Dr S. BONY-LENA
Laboratoire de Météorologie Dynamique (LMD/IPSL)
CNRS/UPMC
Tour 45-55, Case 99
4 place Jussieu
F-75252 Paris Cedex 05
FRANCE
Tel: +33-1-44-27-50-14
Fax: +33-1-44-27-62-72
Email: bony@lmd.jussieu.fr

Dr P. BRACONNOT
Laboratoire des Sciences du Climat et de l'Environnement
Unité mixte CEA-CNRS
Orme des Merisiers, Bat. 712, CE-SACLAY
F-91191 Gif-sur-Yvette Cedex
FRANCE
Tel: +33-1-69-08-77-21
Fax: +33-1-69-08-77-16
Email: Pascale.Braconnot@cea.fr

Dr G. FLATO
Canadian Centre for Climate Modelling and Analysis
Meteorological Service of Canada (MSC)
University of Victoria
P.O. Box 1700
Victoria, BC V8W 2Y2
CANADA
Tel: +1-250-363-8223
Fax: +1-250-363-8247
Email: greg.flato@ec.gc.ca

Dr M. GIORGETTA
Max Planck Institute for Meteorology
Bundesstrasse 53
D-20146 Hamburg
GERMANY
Tel: +49-40-41173-358
Fax: +49-40-41173-298
Email: giorgetta@dkrz.de

Dr F. GIORGI
International Centre for Theoretical Physics (ICTP)
Physics of Weather and Climate Section
P.O. BOX 586, Strada Costiera 11
I-34100 Trieste
ITALY
Tel: +39-040-2240-425
Fax: +39-040-2240-449
Email: giorgi@ictp.trieste.it

Dr A. HIRST
CSIRO Atmospheric Research
PMB 1
Aspendale, Victoria 3195
AUSTRALIA
Tel: +61-3-9239-4531
Fax: +61-3-9239-4444
Email: tony.hirst@csiro.au

Professor M. KIMOTO
University of Tokyo
Centre for Climate System Research (CCSR)
5-1-5 Kashiwanoha, Kashiwa-Shi
Chiba 277-8568
JAPAN
Tel: +81-04-7136-4385
Fax: +81-04-7136-4375
Email: kimoto@ccsr.u-tokyo.ac.jp

Dr C. LE QUÉRÉ
University of East Anglia/ British Antarctic Survey
School of Environmental Sciences
Norwich NR4 7TJ
UNITED KINGDOM
Tel: +44-1603-592840
Email: c.lequere@uea.ac.uk

Professor Dr N. NAKICENOVIC
Professor of Energy Economics | Leader Transitions to New Technologies
Power Systems and Energy Economics | and Greenhouse Gas Initiative
Vienna University of Technology | International Institute for
Gusshausstrasse 25-29/37320 | Applied Systems Analysis (IIASA)
A-1040 Vienna, Austria | A-2361 Laxenburg, Austria
Tel: +43 1 58801 37350 | Tel: +43 2236 807 411
Fax: +43 1 58801 37397 | Fax: +43 2236 807 488
Email: naki@eeg.tuwien.ac.at | Email: naki@iiasa.ac.at

Experts/Observers

Dr D. BADER
Director, PCMDI
Lawrence Livermore National Laboratory
P.O. Box 808, L-103
Livermore, CA 94550
USA
Tel: +1-925-422-4843
Fax: +1-925-422-7675
Email: bader2@llnl.gov

Dr H. BANKS
The Met Office
Hadley Centre
Fitzroy Road
Exeter, Devon EX1 3PB
UK
Email: helene.banks@metoffice.gov.uk

Dr C. COVEY
PCMDI
Lawrence Livermore National Laboratory
P.O. Box 808, L-103
Livermore, CA 94550
USA
Tel: +1-925-422-1828
Fax: +1-925-422-7675
Email: covey1@llnl.gov

Dr V. EYRING
Deutsches Zentrum fuer Luft- und Raumfahrt (DLR)
DLR-Institut fuer Physik der Atmosphaere, Oberpfaffenhofen
D-82230 Wessling
GERMANY
Tel: +49-8153-28-2533
Fax: +49-8153-28-1841
Email: veronika.eyring@dlr.de

Dr P. FRIEDLINGSTEIN
LSCE, CEA-CNRS
Bâtiment 712
F-91191 Gif-sur-Yvette
FRANCE
Tel: +33-1-69-08-87-30
Fax: +33-1-69-08-30-73
Email: Pierre.Friedlingstein@cea.fr

DR L. GATES
Email: wlgates@gmx.net

Dr J. GREGORY
Met Office
FitzRoy Road
Exeter, Devon EX1 3PB
UK
Tel: +44 1392 884542
Email: jonathan.gregory@metoffice.gov.uk

Dr G. HEGERL
University of Edinburgh
School of GeoSciences
Email: Gabi.Hegerl@ed.ac.uk

Dr J. MURPHY
Head of Climate Prediction
Met Office
Hadley Centre
FitzRoy Road
Exeter EX1 3PB
UK
Tel: +44 1392 886658
Fax: +44 1392 885681
Email: James.murphy@metoffice.gov.uk

Dr A. PIRANI
International CLIVAR AOS Program
Princeton University
Forrestal Campus
Princeton, NJ-08544
USA
Tel: +1-609-258-4183
Fax: +1-609-258-2850
Email: anna.pirani@noc.soton.ac.uk

Dr T. STOCKDALE
ECMWF
Seasonal Forecasting
Shinfield Park
Reading, Berkshire RG2 9AX
UK
Tel: +44-1189-499-117
Fax: +44-1189-869-450
Email: Tim.Stockdale@ecmwf.int

Dr K. TAYLOR
Lawrence Livermore National Laboratory
L-103
P.O. Box 808
Livermore, CA 94550
USA
Tel: +1-925-423-3623
Fax: +1-925-422-7675
Email: taylor13@llnl.gov

Dr G. TSELIOUDIS
Research Scientist
NASA/GISS, Columbia University
Department of Applied Physics
2880 Broadway
New York, NY 10025
Tel: +1-212-678-5578
Fax: +1-212-678-5552
Email: gtselioudis@giss.nasa.gov

Dr D. VERSEGHY
Meteorological Service of Canada
Climate Research Branch
4905 Dufferin Street
Downsview, Ontario M3H 5T4
CANADA
Email: Diana.Verseghy@ec.gc.ca

WCRP

Dr A. HENDERSON-SELLERS
Director, World Climate Research Programme
c/o WMO, 7bis Avenue de la Paix
Case Postale No 2300
CH-1211 Geneva
SWITZERLAND
Tel: +41-22-730-8246
Fax: +41-22-730-8036
Email: AHenderson-Sellers@wmo.int

Dr V. SATYAN
Director, World Climate Research Modelling
World Climate Research Programme
c/o WMO, 7bis Avenue de la Paix
Case Postale No 2300
CH-1211 Geneva
SWITZERLAND
Tel: +41-22-730-8418
Fax: +41-22-730-8036
Email: VSatyan@wmo.int

APPENDIX III

A proposal for coordinated experimentation to study multi-decadal prediction and near-term climate change

Draft 5th October 2007

Proposal elaborated by a WGCM/WGSIP/CLIVAR/WCRP sub-group:

Tim Stockdale, Gabi Hegerl , Jerry Meehl, James Murphy, Ron Stouffer, Marco Giorgetta, Masihide Kimoto, Tim Palmer, Wilco Hazeleger, Detlef Stammer, Ben Kirtman and George Boer.

Aims: This proposal describes coordinated experimentation to enable:

1. estimation of the expected climate for the period 2005-2030, relative to recent climate, together with credible uncertainty ranges.
2. study of the errors, uncertainties *[and processes?]* in multi-decadal predictions

In particular:

- the role of initial conditions, and the methods by which they can be specified
- the value of higher resolution models
- sensitivity to choice of model, and effectiveness of multi-model composites
- testing actual level of error in decadal predictions against that expected from ensemble spread and estimated initial condition uncertainty
- uncertainty arising from unknown future forcing (tropospheric aerosols, volcanism)
- comparison of predictions based on different techniques

The final analysis of expected climate in 2005-2030 will:

- Aim to give guidance on the changing risk of extremes
- Aim to give guidance on the possibility of changes in the monsoons
- Aim to arrive at probabilistic forecasts of future climate in this period using a variety of methods and sources, including multi-decadal model predictions, empirical predictions based on attribution of observed changes to date, and results of other modeling studies.

This proposal assumes that complementary coordinated experimentation will take place both to assess the longer term response of the climate system to anthropogenic forcing and possible stabilization scenarios; and to study the evolution of climate from pre-industrial times to the present day, including issues of detection and attribution. Note that the reference integrations called for in this proposal may be used to augment the study of observed climate change in the late 20th century, but are not on their own sufficient to characterize the recent past. It is hoped that the design of "20th century" experiments and the reference experiments described here will dovetail together.

Baseline modeling requirement:

- 10 member ensemble, 25-yr integrations, start dates 1st November 1960, 1980 and 2005. (Actual integration length should be 25 years and 2 months).
- Started from observationally-based initial conditions/analyses, either as anomalies or full fields.
- All forcings should be included as observed values for past dates, together with the best estimate for the future. CO_2 is prescribed, not determined by carbon cycle models.
- A single GHG scenario, namely that to be used in the longer AR5 integrations which leads to a stabilization of 4.5 W/m2 radiative forcing. (Available mid-2008).
- For models without active chemistry, specified future ozone as developed for the AR5 runs. *[For models with chemistry: should past ozone be calculated by the model, or prescribed from observations - ie is ozone a forcing or a response?]*
- History of volcanic aerosol should be included in runs from past dates, and baseline runs should assume no volcanic eruptions in future.

- Where possible, this protocol should describe the forcing to be used.
- *[Needed: more detail on the exact specification of the forcings which are recommended to be used, the tolerances that will be allowed, and on the requirement for adequately documenting what in fact was used]*

Reference integrations:

- Enhancement of a standard ensemble of 20th century transient runs to 10 members, starting on 1 Jan 1960, and running to the end of 2030.
- Initial conditions from standard full-length 20th century transient runs, with perturbations, use of lagging, linear combinations or other methods to create the necessary ensemble size.
- Forcings as in prediction runs: all observed forcings for past, then switch to a standard scenario with no future volcanoes up to 2030.
- Ensemble of reference integrations are requested from as many models as possible. They (i) provide a "control" against which the initialized runs can be compared; (ii) provide a well sampled and reasonable length dataset for detection and impacts work; and (iii) provide some sort of 'back-up' for assessing likely climate change up to 2030, should the initialized runs prove to be problematic to analyze. Models which are unable to provide reference runs will still be accepted in the main experiment(s).

Extensions:

- Runs from a bigger set of start dates, including in the first instance 1965, 1970, ….2000. (These dates coincide with those used in the ENSEMBLES project).
- Multiple resolution runs. Where the 25 year runs are made with a substantially higher resolution than is used for standard historical and scenario runs, it is encouraged to repeat the 25 year runs described here with the standard resolution model, to allow assessment of resolution dependence. It is expected that most of the scientific exploration of initialization method and other sensitivities will be done at a modest resolution, even if the final projections to 2030 (and corresponding reference runs from 1980 or earlier) are repeated at high resolution.
- Experiments 'inserting' externally produced ocean initial condition anomalies into coupled models *[Some groups will be working on this anyway - is there a benefit to coordinating experimentation here? Eg cross-transplants of anomalies between different model systems to assess relative loss of skill]*
- Swapping ocean initial conditions between a pair of years, eg chosen to reflect different states of the MOC, as a sensitivity experiment *[Which years would be best to choose?]*
- Removal (from past start dates) or addition (to runs to 2030) of volcanic aerosol, to assess forecast sensitivity to eruptions (and model differences in this). For past dates, repeating the 1960 and 1980 starts with only the decay of the initial volcanic aerosol load would be a good sensitivity experiment. For the 2005 start, we might add a suitably defined "standard volcano" eg in 2010.
- Small ensemble runs with evolving chemistry and/or alternative tropospheric aerosol scenarios

Timelines, issues and interaction with other coordinated experimental projects:

- Early experimentation to test initialization methods is desired. The ENSEMBLES project is running decadal prediction experiments with start dates 1960, 1965, ... 2005 (but with volcanic aerosol turned off). Some models will only run 10 year integrations, other models will run 30 years. ENSEMBLES integrations are scheduled to complete in August 2008, and the project finishes in August 2009. They will not follow the detailed protocol here (eg choice of GHG scenario), but can be learnt from. *[Are there any other groups that would like to coordinate early experimentation with ENSEMBLES? Would this be helpful or not?]*
- High resolution runs. Because these are expensive, it would be better to wait until information is available on the impact of choice of initialization. On the other hand, some groups at least may not be able to delay the start of them beyond 2009. We hope that observationally based initial conditions will be (robustly) better than those from 20th Century transient runs, but this has yet to be established. Having integrations using both methods would be helpful in deriving predictions based on a range of

techniques, particularly with the goal of fully exploring uncertainty. *[Are groups planning high resolution runs able to run both the initialized and reference integrations at high resolution? How would reference integrations be initialized without a high resolution run from pre-industrial climate? Could such a run be afforded? Or will high resolution runs be mostly restricted to just the baseline experiment because of cost?]*

- Initialization and forecast of sea-ice may be an issue. Loss of Arctic summer ice cover may be accelerating, and little or no experience exists on the initialization and prediction of this.
- The output of the model integrations needs to be defined. *[This is a substantial task. We presume that a suitable starting point will be the basic output requested for the longer climate change runs. But this may need additions (or possibly subtractions) to meet the needs of this particular experimentation, eg in characterizing extremes, and to be feasible for the high resolution runs.]*
- Data handling needs to be defined. It is hoped *[but needs to be confirmed]* that data from the central parts of this proposal can be archived as part of the general AR5 archive, to allow access by the appropriate part of the scientific community. Some of the additional experimentation may need to be handled separately by the groups involved.
- The protocol described here calls for 25 year long runs, with the last integration starting in November 2005 and ending in 2030. *[Is 25 years the right length (compared to eg 30 years?). Is 2005 a suitable "end" start date for the coordinated experimentation, bearing in mind that in some cases real-time initial conditions may not be available? Will groups undertake regular (annual? 5 yearly?) updates of their short-term climate runs anyway, once the background runs have been established?]*
- The protocol described here calls for 10 member ensembles, although it envisages that some additional runs eg with interactive chemistry might use smaller ensembles. Larger ensembles have two big benefits: they allow better visibility of changing climate and initial condition decadal signals against the relatively noisy background of short climate runs; and they allow decent sampling of "extreme" events: a ten member ensemble gives 100 model years of integration in a single calendar decade. *[Have we got the balance right between cost and benefit in asking for 10 members? Is 10 members an absolute requirement or an aspiration? Do we apply a cut-off, and if so, what should it be? Can we accept smaller ensembles (eg 3) for the extensions? ENSEMBLES specify a minimum size of 3, for a larger set of start dates and no focus on extreme events]*
- The final document should contain some discussion/guidance on possible initialization strategies. *[Contributions for this are welcome.]*
- Models can also be tested in seasonal forecast mode following the TFSP protocol. (Start dates of 1 Nov. are consistent with this). Compatibility of the seasonal and multi-decadal runs is encouraged where appropriate.
- Other groups may want to propose coordinated experimentation looking at particular processes relevant to multi-decadal forecasts. Compatibility between such coordinated experimentation and this proposal is encouraged where appropriate.
- As mentioned above, it is important to make sure that the reference integrations mentioned here fit into a proper general framework for 20th century climate.

Maximizing the Value of Community-Coordinated Experiments with AOGCM[1]s and ESM[2]s

Karl E. Taylor
15 August 2007

1. Introduction.

Hibbard, et al. (2007) recently proposed climate model simulations designed to explore the possible consequences of human activities on 21st century climate. Amongst the communities who would make use of these simulations would be those considering climate change's impact on ecosystems and society and its implications for energy policy. The proposed strategy originated at an Aspen Global Change Institute in 2006 and will be subsequently referred to as the "Aspen strategy." Here I propose additional simulations that are essential if we are to understand and interpret the range of model responses that can be expected in the so-called "long-term (2005-2100 and beyond)" experiments of the Aspen strategy. I also argue that some simulations included in the Aspen strategy can be eliminated if the new simulations are accepted, so that the revised strategy suggested here will be no more burdensome than the original in terms of resources required of the climate modeling groups.

Coordinated "standard" climate model experiments, like those proposed by Hibbard, et al., allow us to establish a lower bound on uncertainty in projected changes, since the models (if equally reliable) give a range of responses that must reflect limits of our understanding. All of the models may of course have common errors, which means that the true uncertainty in projections is unknown, but the true trajectory of the climate system could certainly lie outside the range of model trajectories. Although there is interest in learning the degree to which models disagree in their projections, there is considerably more value in determining the *reasons* for the spread of responses. As part of the Aspen strategy, the "long-term" experiments were designed to determine the importance of carbon cycle feedbacks. It is unfortunate, however, that under the Aspen strategy, carbon feedbacks alone would be diagnosed, while other equally uncertain and perhaps even more important feedbacks would remain unexamined. The motivation here for proposing additional "idealized" experiments is to round out the suite of experiments, so that a truly comprehensive analysis of all the important feedbacks can be undertaken.

2. Additional experiments.

The issue addressed first is what experiments are needed to diagnose "traditional" climate feedbacks involving processes such as clouds, water vapor, and sea ice (in addition to carbon cycle feedback, which can already be partially assessed under the original experiment design). The first step to quantifying these feedbacks is to determine the so-called "radiative forcing." Radiative forcing is the immediate (or "fast") response of the climate system (as gauged by changes in the net radiative flux at the top of the atmosphere) to some imposed change (e.g., an increase in aerosol or greenhouse gas concentrations). In the experiments proposed under the Aspen strategy, estimating the radiative forcing as it evolves for each scenario will be difficult because aerosols will be included and in contrast to carbon dioxide, their radiative impact does not follow a simple scaling relationship to their global mean concentrations.

---

[1] "AOGCM" stands for coupled "atmosphere-ocean general circulation model" or "atmosphere-ocean global climate model."

[2] "ESM" stands for "earth system model," which here means an AOGCM coupled to a carbon cycle model.

The new experiments proposed below differ from the arguably more “realistic” Aspen strategy scenarios in two important ways:

1. Increases in atmospheric CO2 concentration will be prescribed, while aerosols and other greenhouse gases will be held fixed. It would be open to community discussion whether the CO2 concentration should increase in an idealized way (e.g., 1%/year increase) or should evolve more realistically[3]. This approach of specifying a single, well understood forcing agent facilitates diagnosis of the evolving radiative forcing and enables a full feedback analysis.

2. The simulations, because they are idealized, can be forced by a somewhat exaggerated rate of increase of atmospheric CO2 concentration, which enhances signal to noise and reduces the length of simulations (or the size of ensembles of simulations) required to obtain statistically robust results. In particular, during most of the 20th century segments of the Aspen strategy scenarios, there will little value in evaluating carbon cycle feedbacks because the changes in CO2 concentration and climate are relatively small during this period (compared to the 21st century). In the additional experiments proposed here we can immediately impose a larger rate of change, perhaps characteristic of the late 20th century or even exaggerated somewhat.

In addition to a control run, the simulations needed for a comprehensive feedback analysis are listed in Table 1. These simulations will hereafter be referred to as “CO2-only experiments.” The first three experiments are multi-decade simulations to be carried out with ESMs, but Expt. A can also be performed by models without carbon cycle components. For ESMs these experiments will allow us to assess the strength of carbon cycle feedbacks following methodology similar to that applied in Friedlingstein et al. (2003). Expts. D and E are essential for determining the “fast” responses[4] of the carbon and climate systems, and they enable a full diagnosis of all the important feedbacks. [Expts. D and E are identical for models without carbon cycle components.] Further discussion of the diagnostic strategy enabled by this suite of experiments is found in section 3; suffice it to say here that unlike the original Aspen strategy, it will be possible to assess reasons for differences among climate model projections that go well beyond considering the differences in their carbon cycle responses. Moreover, in the case of carbon response, the new set of experiments will provide a more complete breakdown into the various elements that constitute the total carbon feedback.

---

[3] A more realistic prescription might be to specify a CO_2 concentration that yields a global mean radiative forcing equal to the combined net forcing attributable to changes in the full suite of time-varying atmospheric constituents (according to some agreed upon scenario). This "equivalent CO_2" concentration would be "seen" by the model radiation codes, but the carbon cycle would see only that portion of the *equivalent* CO_2 attributable to CO_2 itself.

[4] The term “fast response” is used here to refer to earth system model responses that occur on time-scales much shorter than the response time of the oceans. Included among “fast responses” are so-called “radiative forcing” and “stratospheric adjustment.” See section 4 for further discussion.

Table 1: Summary of experiments with forcing by CO2 only.

Expt.	model type†	priority	prescribe	radiation sees	carbon cycle sees	length of run
A	All	high	—	evolving CO2	evolving CO2	many decades
B	CCC-only	high	—	evolving CO2	control CO2	many decades
C	CCC-only	medium	—	control CO2	evolving CO2	many decades
D	All*	high	From control: ocean surface conditions‡ and distribution of vegetation types	doubled CO2	doubled CO2	several years
E	CCC-only*	high	From control: ocean surface conditions‡ and distribution of vegetation types	doubled CO2	control CO2	several years
F	Only* CCCs with dynamic vegetation	medium	From control: ocean surface conditions‡ From Expt. A at time of doubled CO2: distribution of vegetation types	doubled CO2	control CO2	several years

† "All" means AOGCMs with and without carbon cycle components. "CCC-only" means "coupled carbon climate models only."
** The ocean and sea-ice components are inactive in D, E, and F since climatological monthly-mean ocean surface conditions are prescribed in these experiments.*
‡ sea surface temperature, sea ice, and the partial pressure of dissolved CO2 at the ocean surface.

Expts. A and C are similar to the first two experiments proposed in Hibbard, et al. (2007; see their figures 1 and 2), but without the complications of evolving aerosols or atmospheric chemistry. The similarity of the original experiments to the CO2-only set proposed here raises the question of whether all of the original experiments are really necessary. I would argue that the second Aspen strategy experiment becomes largely superfluous because its sole purpose is to determine the strength of carbon feedback induced by climate change itself. The same information (and more) can be diagnosed with the CO2-only experiments, so the original experiment 2 can be eliminated.

Perhaps less obvious is the fact that the third experiment of the Aspen strategy is also of little additional value. Once the carbon cycle feedbacks and the feedbacks determining global climate sensitivity have been diagnosed in the CO2-only experiments, we should be able to estimate with reasonable accuracy the outcome of experiment 3, relying on EMIC[5]s.

Under the Aspen strategy there are good reasons to consider more than one concentration scenario for greenhouse gases and aerosols (say, both a high and a low emissions scenario), so that at a minimum, we should expect that six multi-decadal simulations would be required (assuming all three of the original experiments are performed for each scenario). Under the new strategy proposed here, the number of multi-decadal simulations would overall be reduced to five, since only experiment 1 of the Aspen strategy would be retained (accounting for two simulations – one for each scenario) and only three CO2-only simulations would be added. Moreover, the multi-

[5] "EMIC" stands for "earth-system model of intermediate complexity."

decadal idealized runs would be shorter than those proposed under the Aspen strategy because the initial rate of CO2 increase would be greater.

3. Further description of the experiment strategy.

The main advantage of the suite of ESM experiments proposed here is that the reasons for differences in model responses can be identified. The design enables a unified analysis of feedbacks in the models (see Taylor, 2007) that would be impossible under the Aspen strategy[6]. The analysis will build on the Hansen et al. (2005) method of estimating the radiative "forcing," extending it to include the carbon cycle's "fast" responses to imposed changes. The feedback analysis will rely on a number of established techniques (e.g., Soden and Held, 2005; Taylor et al., 2007), and will yield information concerning both the global mean feedbacks and their regional strength, following the Boer and Yu (2003) approach. The recent work by Forster and Taylor (2007) can also be drawn on to help diagnose the equilibrium climate sensitivity from the transient response to a forcing that is changing with time. All of these analyses are based on monthly mean fields that are routinely saved when running climate models.

Experiment 1 of the Aspen strategy provides information of use to those studying the impacts of projected climate change on ecosystems and society, along with economists and energy policy makers who can explore options for meeting different targets of CO2 emissions. A minimum of two different scenarios (of prescribed greenhouse gas concentrations) will likely be considered, selected to bracket the range of target emissions that might be achievable. There is considerable value in including all the important forcing agents in these experiments (including the various greenhouse gases and aerosols). The complexity of the mix of forcings that need to be considered, however, will make it difficult to understand why models, when forced similarly, yield different projections. Moreover, unless the additional experiments proposed here (A-E) are performed, we will find ourselves still unable to explain why the range of model projections is as large (or as small) as it is.

Each of the CO2-only experiments (A-E, defined here) will serve several purposes, but, summarized below, is their relevance for understanding differences in model projections:

Expt. A: This experiment, spawned from the control run, is similar to experiment 1 of the Aspen strategy, but with changes prescribed only for CO2 concentration (i.e., with aerosols and other greenhouse gases held fixed). The differences in response among models performing this experiment should be similar in many ways to their differences under the less-idealized experiment 1 of the Aspen strategy.

Expt. B: Conditions identical to Expt. A are imposed, except that the land and ocean surfaces "see" atmospheric CO2 concentration fixed at the control value. Importantly, this simulation will be spawned from the same point in the control run as Expt. A. The difference between this experiment and Expt. A isolates the carbon cycle changes due to climate change alone (without the additional changes, like the fertilization effect, induced by increasing CO2 concentration).

Expt. C: Conditions identical to Expt. A are imposed, except that to the atmospheric radiation code, the atmospheric CO2 concentration appears to be held fixed at its control value. Importantly, this simulation will be spawned from the same point in the control run as Expt. A. This experiment is similar to experiment 2 of the Aspen strategy. The

[6] By "unified analysis" I specifically mean that both the carbon cycle feedbacks and other climate feedbacks can be assessed in a common way so that, for example, the spread of feedbacks associated with clouds can be directly compared with carbon cycle feedbacks

difference between this experiment and Expt. A isolates the carbon cycle changes due to increasing CO2 concentration alone (under conditions of minimal climate change). The difference between the sum of responses in Expts. B and C and the response in Expt. A must be due to nonlinearities. If we assume linearity, then Expt. C can be considered low priority.

Expt. D: This experiment with SST's prescribed from the "control" climatology (along with sea ice, the partial pressure of dissolved CO2, and the distribution of vegetation types) serves to isolate the ("fast") responses of the climate and carbon systems to doubled CO2 in the absence of climate changes governed by the time-scales of the ocean. Although changes in vegetation types (dynamic vegetation) are suppressed, other ecosystem responses to the increased CO2 concentration (e.g., changes in stomatal resistance and leaf area index) are active. The evaluation of "radiative forcing" from this experiment follows the method proposed by Hansen, et al. (2005). Surface fluxes of carbon would also be monitored.

Expt. E: Conditions identical to Expt. D are imposed, except that the land and ocean surfaces "see" atmospheric CO2 concentration fixed at the control value. This experiment isolates the radiative forcing (and other "fast" climate responses) attributable to the direct impact of CO2 increases on atmospheric radiation (with the responses of the carbon cycle largely suppressed).

Expt. F: Conditions identical to Expt. E are imposed, except that the distribution of vegetation types is prescribed to be identical to that found in Expt. A at the time of doubled atmospheric CO2 concentration. The difference between the terrestrial carbon fluxes found in Expt. E and in the control run provides a measure of the importance of the changes in distribution of vegetation on the carbon cycle in the absence of appreciable climate change.

4. Discussion and remaining issues.

The "long-term" experiments proposed in the Aspen strategy attempt to serve the needs of the IPCC Working Groups 2 and 3, but are insufficient for explaining the range of projections that surely must be expected. Thus, these experiments are not optimally conceived and fail to meet the needs of Working Group 1 (WG1). The alternative set of experiments proposed here retains the main experiment of value to WG's 2 and 3 (i.e., experiment 1 of the Aspen strategy), but augments it with slightly more idealized experiments (Expts. A-C) and three short diagnostic simulations (Expts. D-F) that will allow a full and unified feedback analysis of the models to identify factors responsible for their differences. Computationally, the revised suite of experiments will be no more burdensome than the original.

As indicated before, in diagnosing ESMs it is useful to distinguish between "fast" and "slow" responses to rising CO2 concentrations. "Fast" responses include responses that are a direct consequence of the change in CO2 concentration (e.g., enhancement of the "greenhouse" effect), but also other responses occurring on time-scales less than a few months (e.g., "stratospheric adjustment", stomatal response in plants). In the case of imposed changes in radiatively active constituents like CO2, "fast" responses are sometimes referred to as "radiative forcing," but this term is not generally appropriate. For example, if we were to impose regional increases in soil moisture to mimic the effects of irrigation, this would have little direct effect on radiation but would immediately affect evaporation. The evaporation might lead to changes in atmospheric water vapor which could substantially impact radiation (among other things), resulting in climate change. The key is that there is a "fast" response of the system (without necessarily any change in global temperature), which affects the earth's radiative balance.

Examples of "fast" responses to increasing CO2 include
- a reduction in outgoing longwave radiation (i.e., an enhanced "greenhouse" effect),
- thermal response of the stratosphere to altered radiation (i.e., "stratospheric adjustment"),
- enhanced flux of CO2 into the ocean due to increased difference in partial pressures,
- a response of plant stomata that tends to reduce evapo-transpiration,
- an increase in net primary productivity in plants,
- a possible cloud component induced by "fast" tropospheric adjustment (Gregory and Webb, 2007)
- Fast responses of the land surface.

"Slow" responses important on longer time-scales have often been termed feedback responses, and these include
- changes in the types of vegetation found in natural ecosystems (represented in dynamic vegetation models),
- changes in ocean circulation and temperature that affect its ability to absorb CO2,
- a multitude of feedback processes traditionally found in AOGCMs, which are in most cases roughly proportional to the magnitude of global temperature change (e.g., clouds, water vapor, snow and sea ice) and therefore evolve on the same time-scale as global temperature.

The traditional separation of "feedback" from "radiative forcing" has proved useful, but both of these are in fact responses of the climate system to some imposed perturbation. What distinguishes the two is that feedbacks, as traditionally defined, only become evident if the ocean temperature changes (which occurs relatively slowly), whereas "radiative forcing" is a "fast" response to the perturbation. It should be stressed that both responses contribute to the change in global temperature.

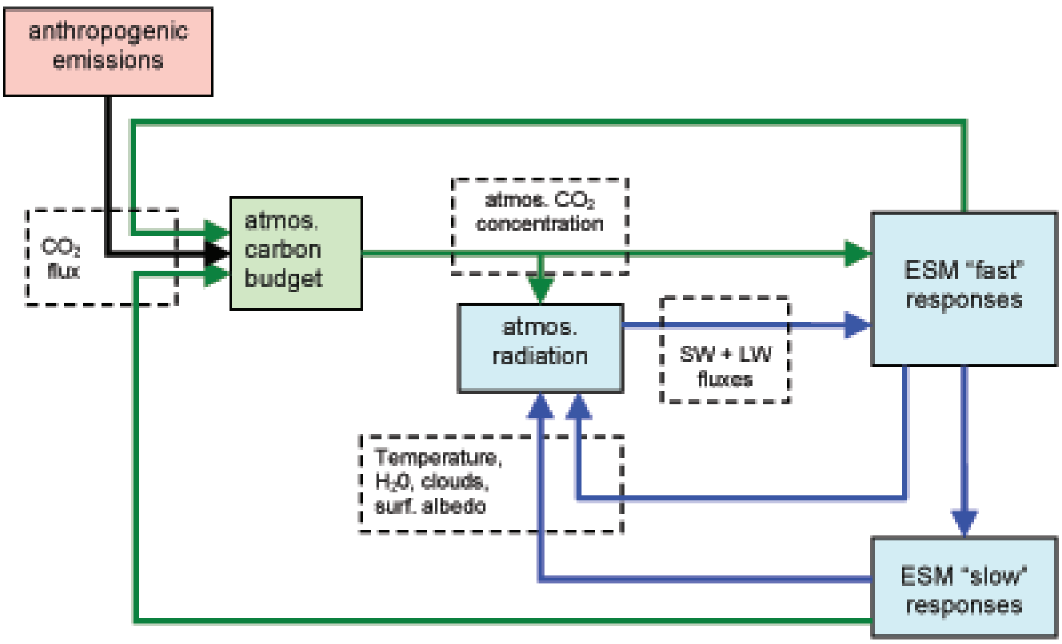


Figure 1: Carbon cycle and climate feedback loops. The blue arrows indicate feedback responses that are active, even in the absence of a carbon cycle. The green arrows indicate carbon cycle feedback loops.

Figure 1 shows schematically how the carbon cycle and climate are linked, with emphasis on important feedback loops. As discussed above, the diagram distinguishes between "fast" and "slow" responses. The blue arrows indicate feedback responses that are active, even in the absence of a carbon cycle. The green arrows indicate carbon cycle feedback loops. If we can quantify the strength of the various feedback responses of the system, then progress in understanding the system may be accelerated by focusing on those processes that are most important (and most uncertain) in determining overall climate sensitivity. If the range of model estimates of a given feedback is large, further work is clearly warranted (although the converse is not necessarily true).

The suite of experiments proposed here will make it possible to characterize across all models the relative importance of the following feedbacks:

- surface albedo feedback, separately accounting for
  - sea ice feedback
  - snow feedback
  - vegetation and soil albedo feedbacks
- shortwave cloud feedback
- clear-sky longwave feedback (associated primarily with water vapor and lapse rate changes)
- cloud effect on longwave feedback

For carbon cycle feedbacks, methods like those described by Friedlingstein et al. (2003) will be supplemented with the information obtained from Expts. E and F, which allows us to assess the "fast" responses of the surface carbon fluxes, yielding a more comprehensive understanding of:

- marine carbon cycle responses, separately accounting for
  - "fast" responses to rising CO2, such as ocean carbon uptake in the absence of any change in the ocean carbon reservoir.
  - ocean carbon uptake (in the absence of climate change, but accounting for slow changes in ocean acidity)
  - the effect of climate change on ocean carbon uptake
- terrestrial carbon feedbacks, separately accounting for
  - "fast" responses to rising CO2 in an unchanging climate with the biomass, vegetation types, and soil carbon essentially unchanging (e.g., CO2 fertilization impact on stomatal regulation of evapotranspiration)
  - impact of rising CO2 on carbon fluxes between the atmosphere and terrestrial biosphere (in the absence of climate change, but allowing the biosphere to fully respond and evolve)
  - the dynamic response of vegetation types to enhanced CO2 and its impact on terrestrial carbon fluxes (in models with dynamic vegetation).

Note that in all of the above it will be possible to resolve the variations in these feedbacks both spatially and seasonally. This should facilitate identification of the underlying processes responsible for the differences in feedback strength among models.

There are several remaining issues that will benefit from community input. Among these are the following:

- How many different scenarios should be considered in experiment 1 of the Aspen strategy? It would seem reasonable that at least one high stabilization scenario and one low stabilization scenario should be required, but it is likely that based on these two, the climate consequences of other scenarios could, with the help of EMIC's, be obtained via interpolation. This issue should be central to the discussions at the IPCC Scenarios Consortium "Expert Meeting" planned for September 2007.
- What are the scientific tradeoffs between consideration of additional scenarios versus

carrying out additional "realizations" of each experiment (i.e., ensembles of simulations)? This issue should be raised at the next WGCM meeting.

- How many years are needed to adequately characterize the "fast" responses in Expts. D and E (i.e., how long do the simulations need to be in order to reduce "noise" to acceptable levels)?
- What evolving concentration of CO2 should be prescribed in Expts. A-C? Among the options might be an idealized increase of 1%/yr (consistent with earlier CMIP exercises) or a more "realistic" prescription that would yield a global mean radiative forcing equal to a net forcing reflecting combined changes in the full suite of time-varying atmospheric constituents (according to some agreed upon scenario).

Acknowledgement: I thank David Lobell and Bala Govindasamy for instructive discussions concerning the carbon cycle. I am grateful to Jonathan Gregory and Mark Webb for helpful comments and suggestions on the experiment design. Remarks from Pierre Friedlingstein, Peter Cox, and Olivier Boucher have been helpful in clarifying aspects of the Aspen Strategy and making known some of the needs of WG2 and WG3.

References.

Boer, G.J., and B. Yu, 2003: Climate sensitivity and response. *Climate Dyn.*, **20**, 415-429, DOI 10.1007/s00382-002-0283-3.

Forster, P.M.D., and K.E. Taylor, 2006: Climate forcings and climate sensitivities diagnosed from coupled climate model integrations. *J. Climate*, **19**, 6181-6194.

Friedlingstein, P., J.-L. Dufresne, P.M. Cox, and P. Rayner, 2003: How positive is the feedback between climate change and the carbon cycle? *Tellus*, **55B**, 692-700.

Gregory, J.M., and M.J. Webb, 2007: Tropospheric adjustment induces a cloud component in CO2 forcing. In preparation.

Hansen, J., Mki. Sato, R. Ruedy, L. Nazarenko, A. Lacis, G.A. Schmidt, G. Russell, I. Aleinov, M. Bauer, S. Bauer, N. Bell, B. Cairns, V. Canuto, M. Chandler, Y. Cheng, A. Del Genio, G. Faluvegi, E. Fleming, A. Friend, T. Hall, C. Jackman, M. Kelley, N. Kiang, D. Koch, J. Lean, J. Lerner, K. Lo, S. Menon, R. Miller, P. Minnis, T. Novakov, V. Oinas, Ja. Perlwitz, Ju. Perlwitz, D. Rind, A. Romanou, D. Shindell, P. Stone, S. Sun, N. Tausnev, D. Thresher, B. Wielicki, T. Wong, M. Yao, and S. Zhang 2005: Efficacy of climate forcings. *J. Geophys. Res.* **110**, D18104, doi:10.1029/2005JD005776 .

Hibbard, K.A., and G. A. Meehl, P.M. Cox, and P. Friedlingstein, 2007: A strategy for climate change stabilization experiments. EOS, **88**, 217, 219, and 221.

Soden, B.J., and I.M. Held, 2005: An assessment of climate feedbacks in coupled ocean-atmosphere models. *J. Climate*, **19**, 3354-3360.

Taylor, K.E., 2007: Integrated feedback analysis in earth system models. Presented at the CFMIP - ENSEMBLES Workshop, Paris, France.

Taylor, K.E., M. Crucifix, P. Braconnot, C.D. Hewitt, C. Doutriaux, A.J. Broccoli, J.F.B. Mitchell, and M.J. Webb, 2007: Estimating shortwave radiative forcing and response in climate models. *J. Climate*, **11**, 2530-2543.

National Oceanography Centre, Southampton
University of Southampton Waterfront Campus
European Way, Southampton SO14 3ZH
United Kingdom
Tel: +44 (0) 23 8059 6777
Fax: +44 (0) 23 8059 6204
Email: icpo@noc.soton.ac.uk